Vol. 41, No. 5 May, 1944

Psychological Bulletin

HULL'S PRINCIPLES OF BEHAVIOR

A Special Review*

SIGMUND KOCH
Duke University

I

The experimental procedures of modern psychology show a degree of rigor hardly warranted by the clumsy and rather literary techniques of theory construction in common use. The recent emphasis by students of psychological method on the virtues of such implements of mature natural science theory as hypothetico-deductive procedure and operational definition has done much to improve this situation. Nevertheless, the dominant temper of present psychology is still crudely, if elaborately, empirical; most research is still devoted to the painstaking accumulation of essentially isolated empirical relationships. Professor Clark Hull has long been in the advance guard of those who are trying to supplement the experimental methods of psychology with theoretical procedures of the requisite precision to make empirical results meaningful. The publication of Hull's new book, *Principles of Behavior*, represents the most potent blow thus far struck against this tradition of ***naive*** empiricism in psychology. At the same time it strikes an equally powerful blow at the complementary disease cluster of literary, programmatic, verbal or speculative theorizing independently of—sometimes in spite of—empirical evidence. For both of these reasons, and for others, *Principles of Behavior* is one of the most important books published in psychology during the twentieth century. No psychologist, whatever his affiliation, can afford the luxury of not reading this immensely challenging work.

Ever since 1929, Professor Hull has been elaborating a comprehensive theory of adaptive behavior based on explicitly stated principles largely deriving from conditioning and other objective learning experiments. Using careful hypothetico-deductive methods and proceeding in cautious, step-wise fashion, Hull has progressively stated, revised and sharpened his postulates and patiently deduced key theorem sequences in a long series of publications. Some of these earlier publications—

* Hull, C. L., Principles of behavior, New York: Appleton-Century, 1943. Pp. x+422.

particularly the recent *Mathematico-Deductive Theory of Rote Learning*—serve as models of the most advanced and rigorous theory construction as yet to be found in psychology. Each of these publications, however, is restricted to the presentation of a fragmentary sub-system of the total theory of behavior which Hull and his colleagues have been designing. Although these formulations have often been impeccable in form and suggestive in content, it has always been difficult to evaluate the overall scope and adequacy of the entire system. This difficulty has now been obviated by the *Principles of Behavior*.

The new book presents the first published statement of the *entire* postulate set tentatively believed to be capable of generating the major phenomena of organismic behavior. The postulates are stated in conjunction with a full discussion of the experimental evidence and theoretical requirements which guided their formulation, and the explanatory power of the system is illustrated by the informal derivation of some fifty important theorems. If the earlier publications gave any reason for doubting the fruitfulness of Hull's constructs, *Principles of Behavior* should go far toward dissipating such doubt. Hull's previous work left no room for questioning the methodological maturity of his *general* approach; the present work makes it no longer possible to question the value and richness of his *specific* theoretical assumptions.

Another general aspect of the book is worth mentioning before going on to more detailed analysis. One of the most resistant obstacles to progress in psychology has been the multiplicity of theoretical language systems to which investigators order their data. This has resulted in a degree of semantic confusion and spurious controversy unique to the social sciences. The prospects for unified theory in psychology seem a little brighter when it is realized that most "rival" psychological theories may, in spite of their terminological individuality, be regarded as largely intertranslatable variants of only two theoretical approaches: "S-R" theory and "field" theory. The feeling has long been prevalent that these two classes of theory are basically opposed on many crucial issues. *Principles of Behavior* marks a major step toward the reduction of this apparent opposition. Certain changes in Hull's system—notably the recognition of the hypothesis of "neural interaction"—make it possible to deduce from S-R assumptions many of the phenomena formerly assumed by the field theorists to be compatible only with their own tenets. Advocates of unified theory in psychology have reason to be highly encouraged by this development.

Since most field theories are largely programmatic in form, while Hull's theory, although far more explicit, is certainly incomplete and tentative, it is impossible as yet to estimate finally the extent of agreement. But the present form of Hull's system, as exemplified in the new book, makes it seem probable that many of the primary molar assump-

tions of field theory are deducible as theorems from Hull's *relatively* less molar postulates. As Hull himself says:

> There is much reason to believe that the seeming disagreements among current students of behavior may be largely due to the difference in the degree of the molar at which the several investigators are working. . . . In the end the work of all those who differ only in this sense may find a place in a single systematic structure, the postulates . . . of those working at a more molar level ultimately appearing as theorems of those working at a more molecular level (21).

As already indicated, the substance of the book consists of the statement and elucidation of a tentative postulate set for an objective theory of organismic behavior. The purpose of this postulate set is no less ambitious than to serve as the basis for all the behavior and social sciences. Sixteen postulates are advanced, and special attention is given to the derivation of four "major theorems" pertaining to behavior phenomena of key importance. As the postulates are introduced, their systemic implications are partially explored by the derivation of some fifty additional corollaries. The experimental evidence serving as the inductive source of each postulate is carefully examined, and the inferential operations involved in the formulation of the postulates are, for the most part, clearly specified. It should be stressed that the book does not pretend to be a statement of the entire system. As its title suggests, the book covers merely the foundation principles or postulates of the system. The few theorem sequences that appear are introduced only in order to clarify the significance of key postulates. Since the range of application of the postulate set is presumably as wide as all the social sciences, most of the theorems remain to be deduced. The author looks forward to a long series of investigations which will progressively derive the test theorem sequences relevant to the phenomena of the various social and behavior sciences.

Since Hull's system has attained the stage of quantitative formulation, the task of clear exposition for non-mathematical readers entails formidable difficulties. Hull meets this problem by developing the material verbally and relegating all mathematical formulations to notes at the end of each chapter. In the main this device has succeeded. With the exception of a few fuzzy passages, the bulk of the book should be intelligible to any educated reader with the courage to expose himself to rigorous statement and close knit reasoning. For purposes of clarity, we will confine the present discussion to the verbal level.

Instead of attempting a complete *précis* of this complex work, we will restrict scrutiny to the major postulates and the high spots of the discussion in which they are embedded. After this outline, we will call attention to some of the many ways in which the present statement of the theory has improved on earlier formulations. Finally, we will offer a few constructive criticisms.

II

The book begins with a lucid discussion of the nature of scientific theory. The first chapter presents an analysis of hypothetico-deductive procedure as employed in the natural sciences. Hull stresses the joint importance of the empirical and theoretical components of science, and describes the function of each. Scientific explanation is identified with the *deduction* of theorems expressing a natural event from one or more primitive assumptions or postulates taken in conjunction with the antecedent conditions of that event.

The second chapter discusses the methodological requirements of "objective behavior theory." Here Hull identifies the task of a molar science of behavior as the isolation of "the basic laws according to which various combinations of stimulation, arising from the state of need on the one hand and the state of the environment on the other, bring about the kind of bahavior characteristic of different organisms" (19). Because of the marked incompleteness of present neurophysiological knowledge, Hull very wisely chooses to elaborate his theory on a molar basis. He stresses the valuable point that the development of an adequate neurophysiology of behavior is itself contingent on the *prior* development of an adequate molar or macroscopic theory of behavior. Further, Hull takes pains to indicate that the difference between the molar and the molecular is a relative one. Although "there are degrees of the molar, depending on the coarseness of the ultimate causal . . . units dealt with," he prefers to keep the causal units as fine as possible—"to approach the molecular just as closely as knowledge of (the neurophysiological) substrate permits" (21). This may serve to unconfuse those who have persisted in assuming that Hull is essentially a physiological theorist, or at least a behavior theorist with one foot in physiology.

Hull believes that the adoption of a molar approach entails the use of "intervening variables" (logical constructs) in behavior theory. Intervening variables "represent entities or processes which, if existent, would account for certain events in the observable or molar world" (21). In the physical sciences they take the form of such high order constructs as atoms and molecules, while a concept like "habit" is a typical intervening variable in the behavior sciences. Logical constructs of this sort, however, can only be legitimately used when they are connected by explicit functional relationships with "directly observable events both as antecedents and consequences."

The structural outlines of Hull's system may be viewed as a complex chain of intervening variables interpolated between the independent variables which serve as the observable antecedents of behavior, and the dependent variables which represent measurable

aspects of the consequent behavior itself. Among the chief *independent* variables are stimuli (S), conditions which produce drives (C_D), conditions which eventuate in need reduction (G), etc. The chief *dependent* variables are four measurable indices of response (R): "probability of reaction evocation" (p), reaction latency ($_{s}t_{R}$), resistance to experimental extinction (n) and amplitude (A). The multiple functional relationships between these sets of end terms and the mediating chain of intervening variables are defined by the postulate set.

The initial intervening variable is introduced in *Postulate 1.* This is the "afferent neural impulse" (s) which is generated by the contact of a stimulus energy (S) with a receptor organ. Neurophysiological data suggest that, after a latency period, s rapidly attains maximal intensity and then gradually diminishes to a low value according to a "simple decay function." It is also assumed that s "continues its activity in the . . . nervous system for some seconds after the termination of the stimulus" (47), diminishing to zero as a decay function of its value at the cessation of the stimulus.

The second postulate introduces the critical hypothesis of "neural interaction." This assumes that all afferent impulses simultaneously active in the nervous system interact with each other, partially changing the nature of each. Thus, a new intervening variable—an afferent impulse as modified by neural interaction ($\bar{s}$)—is introduced. This postulate deserves special attention because it introduces into the system for the first time a reference to factors of organizational interdependence of the sort stressed by the field theorists.

The major postulate of the system defines the conditions under which learning occurs. This is *Postulate 4*—the law of "reinforcement" —which introduces the key intervening variable of "habit strength" ($_{s}H_{R}$). Habit strength represents the presumptive neural modification of the organism produced by learning. This modification is conceived behaviorally as an increased tendency for a given receptor activity (s) to evoke a given effector activity (r). On a more molar level, habit strength may be expressed as a relation between the parent stimulus of s (S) and the reaction (R) corresponding to r.

The general conditions which determine an increment of habit strength ($\Delta_{s}H_{r}$) are simply stated: whenever a receptor activity coincides with an effector activity ($_{s}C_{r}$), and is closely followed by the reduction of a need (G), or a stimulus which has been associated with need reduction, there results an increment of habit strength. The *magnitude* of $\Delta_{s}H_{r}$ is a quantitative function of several measurable aspects of the learning situation. The relations of four of these independent variables to the generation of habit strength have been sufficiently investigated to warrant a tentative quantitative statement of

the respective functions. Hull considers each of these relationships along with the relevant experimental evidence in separate chapters. He concludes that $\Delta_s H_r$ is: (1) a positive growth function of the number of reinforcements (N), (2) a positive growth function of the magnitude of need reduction (G), (3) a negative growth function of the delay in reinforcement (t), and (4) a negative growth function of the time difference (t'') between S and R when both are of short duration, or a negative growth function of the time (t') the stimulus acts on the receptor, when S is continuous.

All of these relationships may be expressed in a single complex equation which expresses habit strength as a positive growth function of the number of reinforcements, while the remaining three functions jointly determine the asymptote of the learning curve. Since other factors such as stimulus intensity and vigor of response probably enter into the determination of habit strength, Hull underlines the provisional and incomplete character of the formulation.

Because habit strength is a quantitative construct, a suitable unit of measurement must be introduced. Hull adopts a centigrade unit—the "hab"—which represents "1 per cent of the physiological maximum . . . of habit strength attainable by a standard organism under optimal conditions" (114). Various crude indices of habit strength are available, but the precise operational definition of the hab (and other intervening variable units which we will mention later) is a major unsolved problem of the system. Hull recognizes this (281) and indicates that, at present, units like the hab are used as illustrative devices.

The generality of the law of reinforcement cannot be appreciated without a word about "secondary reinforcement." The occurrence of need reduction (G) constitutes "primary reinforcement." But increments in habit strength may also be mediated by all stimuli which are "closely and consistently associated with G." Such secondary reinforcing agents must be assumed to operate in all learning situations. Experimental evidence suggests that primary reinforcement does not affect stimulus-response coincidences preceding its occurrence by more than about 30 seconds. Thus, the learning of behavior sequences which exceed the range of primary reinforcement is mediated by secondary reinforcement. Recalling that the delay of reinforcement is assumed to generate habit strength according to a *decay* gradient, it is clear that Hull's well known early principle of the "goal gradient" may now be deduced from the gradient of primary reinforcement in conjunction with secondary reinforcement.

Having introduced habit strength, we must now consider further links in the long chain of logical constructs intervening between behavior and its antecedent conditions. *Postulate 5*—the principle of

"stimulus generalization"—introduces the concept of "effective habit strength" (${}_{s}\overline{H}_{R}$). The facts of stimulus equivalence suggest that reinforcement eventuates not only in a connection between the *particular* receptor-effector activities reinforced, but in the connecting of a wide zone of *similar* receptor activities to the reaction in question. Since the actual stimulus of reaction evocation (S) must always differ from the stimulus of original conditioning ($\dot{S}$), behavior cannot be predicted from the habit strength commanded by the learning stimulus. Rather, one must know the effective habit strength commanded by the *evoking* stimulus. Postulate 5 specifies the function linking effective habit strength to habit strength.

Differences between the afferent impulse of original reinforcement and the evoking impulse may be expressed in terms of j.n.d. units (discrimination thresholds) on the corresponding stimulus continuum. This being the case, postulate 5 states that effective habit strength is a joint negative growth function of the strength of the reinforced habit, and the magnitude of the difference (d) between the afferent impulse of original conditioning ($\dot{s}$) and the evoking impulse (s). The form of this function is suggested by Hovland's experiments on pitch and intensity generalization, and the postulate may be formulated mathematically by equations fitted to Hovland's data.

From postulate 4 (reinforcement) and the principle of stimulus generalization, there flows a theorem of considerable importance. This is "Major Corollary I" which maintains that "all effective habit tendencies to a given reaction . . . . which are active at a given time summate according to the positive growth principle exactly as would the reinforcements which would be required to produce each" (199). This principle of "physiological summation" functions in the derivation of many important behavior phenomena. For instance, it accounts for the "paradox" that superthreshold learning is able to occur in spite of the fact that each reinforcement yields a subthreshold habit increment to connections which are never *exactly* repeated. Again, physiological summation facilitates derivation of many properties of "compound conditioned stimuli"—such as the experimentally verified (Woodbury) tendency for a compound to evoke a smaller amplitude of reaction than the sum of the reaction amplitudes evoked by its components when presented separately. To take a final important example, this principle is instrumental in deducing the phenomena of stimulus *patterning*.

Another factor which must obviously be represented among the intervening variables which determine behavior is the *need* state of the organism at the time of action evocation. Needs operate not only in the formation of habits (i.e., in the process of reinforcement), but in "their subsequent functioning" (247). Since this is the case, an inter-

vening variable representing the joint effects of effective habit strength and drive in the causation of behavior is necessary. This requirement is satisfied by the construct of "reaction potential" (${}_sE_R$).

The relationship of ${}_sE_R$ to ${}_sH_R$ and drive intensity (D) is suggested by a very impressive experiment performed by Perin. Using groups of rats in a Skinner box learning situation, Perin determined the resistance of the learned habit to extinction (${}_sE_R$) as a joint function of (1) the number of reinforcements (${}_sH_R$) and (2) the number of hours of food privation (D) preceding extinction. Fitted equations indicate reaction potential to be a multiplicative function of habit strength and strength of drive, at least when these variables are identified with the terms of Perin's experiment.

Perin's experiment plus a survey of much evidence relating to primary motivation suggest two additional postulates. *Postulate 6* assumes that "associated with every drive (D) is a drive stimulus (S_D) whose intensity is an increasing monotonic function of the drive in question" (253). *Postulate 7* states that:

> Any effective habit strength (${}_s\overline{H}_R$) is sensitized into reaction potentiality (${}_sE_R$) by all primary drives active . . . at a given time, the magnitude of this potentiality being a product obtained by multiplying an increasing function of ${}_s\overline{H}_R$ by an increasing function of D (253).

Hull conceives the sensitizing action of drive on habit strength to be mediated by chemical "drive substances" (e.g., endocrine secretions) which are released into the blood stream. In the case of some drives (e.g., sex) these substances presumably initiate the drive stimulus, while in the case of others (e.g., tissue injury) the reverse relation obtains. A programmatic unit for the measurement of drive strength—the "mote"—is defined as one per cent of the maximum possible drive strength for a standard organism. A similar centrigrade unit for the measurement of reaction potential is called the "wat." Postulates 6 and 7 make possible the deduction of many of the major phenomena of primary motivation.

Hull regards experimental extinction as an adaptive mechanism of the first importance because it leads to the elimination of those unfunctional habits which are no longer followed by need reduction. Consequently, some provision for the effects of extinction and related processes must be made among the intervening variables of the system. To this end, Hull introduces the construct of "inhibitory potential" ($\dot{I}_R$), and defines its functional properties in the eighth and ninth postulates.

The analysis of inhibition takes its departure from the Mowrer-Miller hypothesis which assumes, among other things, that whenever a reaction is evoked, there results a condition which acts as a negative

drive state in that it tends to cause a cessation of the activity which produced the condition. This fatigue-like condition is called "reactive inhibition" (I_R). But, if reactive inhibition is tantamount to a negative drive state, then its reduction after each reaction would "constitute a reinforcing state of affairs." This means that all stimuli "closely associated with the cessation of a response" must become conditioned to the corresponding reactive inhibition (I_R), thus resulting in "negative habits" tending toward cessation of the response. In this way there emerges a second inhibitory mechanism: "conditioned inhibition" or ${}_sI_R$. The performance of any reaction necessarily involves the generation of *both* reactive and conditioned inhibition, and it is assumed that the two types of inhibition summate physiologically exactly as would positive habit tendencies. The net result of such summation is called "total inhibitory potential" ($\dot{I}_R$).

Reactive inhibition is implicitly defined in the first part of *Postulate 8*. Succeeding parts of the postulate state two conditions, suggested by available experimental evidence, which govern the amount of *total* or net inhibition ($\dot{I}_R$). The first states that the amount of $\dot{I}_R$ produced by successive reactions is "a simple linear increasing function" of the number of reactions. The second supposes that $\dot{I}_R$ is a positively accelerated increasing function of the *work* involved in the response (Mowrer-Jones experiment). A final portion of postulate 8 assumes that *reactive* inhibition "spontaneously dissipates as a simple negative growth function of time" (300). Postulate 9 defines the notion of conditioned inhibition and indicates that it summates physiologically with reactive inhibition.

Returning to the chain of intervening variables which govern behavior, it is clear that reaction potential (${}_sE_R$) will only determine final reaction to the extent that it overrides total inhibition ($\dot{I}_R$). In order to represent this a new construct—"effective reaction potential" (${}_s\overline{E}_R$)—must be added to the system. Effective reaction potential is defined as reaction potential minus total inhibitory potential. A suggested unit of inhibitory potential—the "pav"—represents "that amount of I_R which will just neutralize one unit of reaction potential" (280).

The introduction of postulates 8 and 9 makes possible the derivation of a large number of familiar learning phenomena. Among the more immediate implications are theorems pertaining to the following phenomena: the generalization of extinction effects, spontaneous recovery, disinhibition, inhibition of reinforcement, reminiscence and the economy of distributed practice. Of great significance is the capacity of these assumptions to generate as a theorem the tendency for organisms (other things equal) to choose behavior sequences involving minimal

energy expenditure. Field theorists have attributed similar phenomena to the operation of a *primary* principle of behavior—variously designated as the law of "parsimony," "least action," etc.

With the construct of ${}_{S}\overline{E}_{R}$ before us, we are still some distance away from the final variable which determines reaction. The next link which must be added to the complex chain of constructs makes allowance for the ubiquitous *variability* of organismic behavior. Even reactions mediated by highly stable habits tend to oscillate in vigor, latency and other ways from one performance to another. *Postulate 10* makes "behavior oscillation" a lawful part of the system by introducing the construct of "${}_{S}O_{R}$" (oscillation of reaction potential).

Learning experiments (Hays and Woodbury, Thorndike), which record the vigor of a great number of reactions performed by individual organisms at the limit of learning, suggest an approximately *normal* distribution of the oscillating reaction intensities. Such data make it reasonable to suppose that "the variability of reaction under seemingly constant conditions is due to the action of an oscillatory force [${}_{S}O_{R}$] upon the effective reaction potential" (313), and that ${}_{S}O_{R}$ "approaches a normal chance distribution" (314). Neurophysiological factors which could conceivably provide the rationale for such an assu mption are the spontaneous variability of reaction latencies and thresholds found in individual axon fibres (Blair and Erlanger), and the random, spontaneous firing of nerve cells (Weiss).

Other properties tentatively assigned to ${}_{S}O_{R}$ in postulate 10 are: (1) it is identified as an "inhibitory potentiality," (2) its "range, maximum and minimum are constant," and (3) the "amount associated with the several habits of an organism at a particular instant is uncorrelated" (319). Since the presence of ${}_{S}O_{R}$ will necessarily depress the value of the associated effective reaction potential, an additional construct is needed to represent the remainder. This is the "momentary effective reaction potential" (${}_{S}\overline{E}'_{R}$).*

This ingenious analysis of variability has important methodological consequences for psychology. It implies that psychological prediction must be essentially statistical in the sense that momentary *individual* behavior can be predicted only within a certain probability range. Again, it suggests that all behavior laws must *necessarily* be formulated in terms of the central tendencies of large numbers of empirical determinations. Finally, Hull believes that, in part, "behavioral oscillation . . . is responsible for the present relatively backward state of the behavior (social) sciences" (319).

An eleventh postulate introduces the notion of the "reaction threshold" (${}_{S}L_{R}$). The obvious fact that in learning situations "several

* A prime (′) by the side of the bar has been substituted for the dot over the bar used in this symbol by Hull.—EDITOR.

reinforcements are frequently required before the stimulus will evoke the reaction" indicates that reaction potential must exceed a certain inertial factor before reaction can be evoked. Thus ${}_sL_R$ may be defined as the minimal effective reaction potential which will evoke observable reaction. *Postulate 11* simply asserts that ${}_s\overline{E}'_R$ "must exceed the reaction threshold . . . before a stimulus will evoke a given reaction" (344).

At this point, the intervening variable chain of the system is completed. Postulates 1 through 11 have quantitatively moored these variables both to their antecedent conditions and to each other. They must now be related with equal explicitness to the various "observable" aspects of the consequent reactions. In other words, the functional relations of the "pivotal" construct ${}_s\overline{E}_R$ to (1) probability of reaction evocation (p), (2) reaction latency (${}_st_R$), (3) resistance to extinction (n) and (4) reaction amplitude (A) must be determined. These relationships are expressed in postulates 12 through 15.

In order to determine the four relationships, Hull analyzes the equations for characteristic empirical learning curves plotted in terms of each of the four reaction measures into two components. One component equation expresses ${}_s\overline{E}_R$ as a function of the number of reinforcements (N), and the second component expresses the particular response measure in question as a function of ${}_s\overline{E}_R$. Since ${}_s\overline{E}_R$ is an unobservable, this is a rather delicate task. Each of the four learning equations, however, turns out to "contain an explicit positive growth function of N." Recalling that the postulates of the system imply that, other things equal, ${}_s\overline{E}_R$ is a positive growth function of N, it is permissible to regard these functions as the equations determining ${}_s\overline{E}_R$. The equations for each of the response measures as a function of ${}_s\overline{E}_R$ may now be derived by replacing the growth expression in each equation by ${}_s\overline{E}_R$.

Using the above procedure, Hull derives the four tentative relationships stated by postulates 12–15. They are: (1) p is a normal probability function of ${}_s\overline{E}_R$, (2) ${}_st_R$ is a negatively accelerated decreasing monotonic function of ${}_s\overline{E}'_R$, (3) n is an increasing linear function of ${}_s\overline{E}_R$, and (4) A is a linear increasing function of ${}_s\overline{E}'_R$.

The final postulate of the system—the sixteenth—deals with the problem of what happens when a stimulus complex gives rise to two or more incompatible reaction potentials. Although learning experimentation suggests the occurrence of "associative inhibition" under such circumstances, Hull is not prepared as yet to state the function determining possible "indirect inhibitory effects." Thus, for the present, Hull merely assumes in postulate 16 that when reaction potentials to two or more incompatible responses are active at the same time, "only the reaction whose momentary effective reaction potential . . . is greatest will be evoked" (344).

After completing the statement of the postulate set, Hull devotes

a very impressive chapter to the derivation of a theorem sequence concerning "the patterning of stimulus compounds." This represents the first detailed quantitative analysis of patterning ever attempted, and so demands more strenuous examination than can be made in this review.

Stimulus patterning means the ability to react to certain combinations or configurations of stimuli "differently than to the component conditions" (376). The principles of the system generate a large number of quantitative theorems which specify the extent to which such patterning will occur under various conditions of learning. The key assumptions which make possible these deductions are the principles of afferent interaction, reinforcement, physiological summation, generalization of reaction potential and generalization of conditioned inhibition. These and other principles of the system imply that, under conditions of differential reinforcement, patterning will tend to occur. In addition, even when a compound is reinforced non-differentially, a limited degree of "spontaneous patterning" will be in evidence. Other theorems indicate that the degree of patterning is an "increasing function of the degree of interaction effects among the components," "simultaneous compounds will be easier to pattern than temporal" compounds, "the longer the time interval separating the components in a temporal compound, the more difficult the patterning" (377), etc.

III

When two psychologists discuss Hull, they are usually talking about different men. Because of the frequent revisions and extensions which have marked the laborious development of the theory, a person's impression of the system often depends on where he discontinued reading the long series of Hull's publications. The plurality of Hulls, distinguishable in successive publications, is certainly confusing, but this hardly justifies evaluation of the system in terms of an outmoded or partial formulation. Many of the difficulties and gaps detectable in earlier statements have been eliminated from the theory in its present form. For this reason, it will be useful to mention some of the many improvements which have been incorporated into the theory as represented in *Principles of Behavior*.

A favorite criticism directed against Hull by field theorists maintains that Hull, like most "associationists," is essentially a *historical* theorist. Presumably, this means that he derives behavior from conditions established in the past history of the organism (habit connections) and neglects the *momentary* determinants of behavior (the "field"). This criticism was never wholly justified, but now is less so than ever. If the field theorists ever had a point it was only to this extent: in certain

early formulations of the theory Hull tended to derive behavior from "excitatory tendencies" directly established by reinforcement. About the only *momentary* "antecedent condition" explicitly taken into account was the stimulus complex (including the drive stimulus) impinging on the organism. This is no longer the case. The only "historical" construct now found in the system is ${}_{S}H_{R}$ (past learning). Hull derives behavior not directly from ${}_{S}H_{R}$, but from ${}_{S}\overline{E}'_{R}$. But ${}_{S}\overline{E}'_{R}$ is a *joint* function of past learning (${}_{S}H_{R}$) and a host of other factors representing *momentary* conditions of behavior such as (1) the position of the stimulus on the generalization gradient, (2) the intensities of all present drive states, and (3) the oscillation of reaction potential.

The point may be clarified by showing the analogy in the functional status of Hull's chief constructs and those of a field theorist like Lewin. Effective habit strength serves the same purpose for Hull that "cognitive structure" does for Lewin: both of these constructs represent the learning organizations which function in the determination of behavior *at a given time*. The calculation of specific values for *both* of these variables *necessarily involves reference to historical factors* (prior reinforcement for Hull, prior field exposure for Lewin). Again, reaction potential is analogous to the Lewinian notion of "psychological force"; in both theories the respective constructs are jointly determined by momentarily active learning structures and concurrent *motivation*.

The only difference that would seem to remain between Hull and field theorists in their regard for momentary behavior determinants is that Hull states their assumed relations to his intervening variables in full and quantitative detail, while the field theorists do not. We must conclude that if due regard for the manifold of momentary factors which determine behavior is what makes a theorist a "field" theorist, Hull is the arch field theorist of them all.

A second change, related in part to the preceding point, concerns Hull's general treatment of motivation. There has always been a certain ambiguity in Hull's approach to the problems of primary motivation. The only functional property formerly assigned to drive (D), in addition to its role as a necessary condition to reinforcement, was its capacity to provide a "persistent source of stimulation" (S_D). Hull appeared to be trying to deduce all of the phenomena of primary motivation from the presence of S_D as a conditionable stimulus during all learning situations (and other non-motivational principles of the system). In *Principles of Behavior*, Hull explicitly assigns an additional property to drive: drive states are associated with a chemical drive substance which exerts a sensitizing effect on the habit structures mediating response evocation. Thus, during the functioning of habits, drives are now seen to have an energizing effect quite independently of their role as con-

ditioned stimulus elements of evoking stimulus compounds. Furthermore, the first attempt on record is made to determine quantitatively the *amount* of this energizing effect on the operation of learned habits, as a function of the strength of the concurrent drive (Perin experiment). In this manner an incomplete, and in some respects inadequate, theory of motivation gives way to one of the most explicit and plausible theories now available.

Two additional improvements of crucial significance must be considered together. The first is the introduction, as intervening variables, of the "afferent neural impulse" (s) and the "efferent discharge" (r). The second is the adoption of the principle of "afferent neural interaction" (postulate 2). Most of the earlier statements of the system phrased all learning connections in terms of stimuli (S) and responses (R) rather than in terms of the *immediate* receptor and effector processes involved. This being the case, the system was forced to approach most of its problems on a much more crudely molar basis than at present, and it was impossible to cope effectively with such basic phenomena as compound conditioning and patterning. Since it is meaningless to speak of mutual interaction effects as occurring among external stimulus processes, the interpolation of s and r was a first necessary step to the treatment of interaction effects of the sort peculiar to psychology. And the next obvious step was the introduction of the principle of neural interaction itself. Only by virtue of these modifications has the system been able to come to grips with configurational problems of the type so vigorously emphasized by the field theorists.

Although it is unrealistic to expect an early *rapprochement* between Hullian "S-R" theory and field theory, these modifications open up a bridge which could conceivably lead to ultimate reconciliation. There is even a possibility for immediate *cooperation*. A basic unsolved problem of Hull's theory is the determination of the conditions which govern the amount and *character* of neural interaction effects. The great mass of field (*gestalt*) experimentation on the organizational determinants of perception may already contain useful hints at some of the answers. Additional work along the same lines may help fill out the picture.

Another advance may be noted in Hull's present analysis of inhibition. In early publications, inhibitory tendencies (inhibitory potential) were assumed to accumulate only as a result of non-reinforcement. The present assumptions about reactive and conditioned inhibition—particularly those suggested by the Mowrer-Miller hypothesis—greatly extend the range of phenomena which the system can adequately explain.

The present form of the system contains so many marked advances

over one or another of the earlier versions that any selection of the *major* improvements is bound to be arbitrary. Those who have not read *Mathematico-Deductive Theory of Rote Learning* will find that the qualitative functions and informal deductions characteristic of the bulk of the early articles have been supplanted by quantitative functions and stricter derivations. Again, at many places (e.g., the delay of reinforcement gradient) quantitative relationships previously assumed have been modified on the basis of fresh experimental evidence. In general, then, the right to make a serious evaluation of Hull's system must be purchased by a careful reading of *Principles of Behavior*.

IV

Now for a few final evaluative comments. It is easy to point to many aspects of the system which are incomplete, based on tenuous experimental foundations, or lacking in rigor. But in most cases this would not be news for Hull. So clearly aware is Hull of the *provisional* character of his formulations, and the presence in the system of many open questions, that one hesitates to raise critical points. Thus, the following remarks are offered as constructively motivated suggestions rather than objections.

One vital question, which demands more explicit discussion than Hull has yet volunteered, surrounds the significance of the empirical constants which appear in the mathematical statement (equations) of many of the postulates. These equations are derived by curve fitting techniques from the empirical data of specific learning experiments. The usefulness of such equations in the derivation of behavior will depend on the *generality* of their parameters. To what extent can it be assumed, for instance, that the growth constants appearing in a given empirical learning curve will hold for (1) all learning situations of the same type, (2) other types of learning situations, (3) other species of organisms in either similar or quite different learning situations? Again, to what extent will individual differences in the values of these constants appear among organisms of the same type in the same learning situation? Is the inability of the system to predict momentary *individual* behavior due to these differences as well as to the oscillation of reaction potential? The reviewer believes that some of these questions can be answered at present. Others must wait on a protracted program of experimental research. Even if *no* final answers are as yet possible, however, such questions should be explicitly raised and at least provisionally discussed.

A second and somewhat related consideration pertains to the principle of primary stimulus generalization (postulate 5). The assumed form of the generalization gradients for quality and intensity is based on

Hovland's experiments on *auditory* (pitch and intensity) generalization. An important but neglected question is whether similar generalization gradients would be found in other sensory modalities. Again, there seems to be a basic ambiguity in Hull's analysis of the stimulus (and afferent impulse) continua along which reinforcement presumably generalizes. A "continuum" into which independently variable stimulus elements (e.g., pitch) may be ordered is one thing. The "continuum" corresponding to a dimension of similarity along which a stimulus *object* may vary (e.g., Hull's example of the rotation of a die on its axis) is another. Generalization effects along these two types of "continua" cannot be treated in the same manner, and will probably not assume the same form.

A third point which suggests itself may actually constitute more a virtue of Hull's procedure than a limitation. This is Hull's tendency to fit equations to empirical curves containing fewer data plots than caution might dictate. Some of the curves to which Hull fits key equations of the system contain no more than three or four data points. For instance, Perin's curve for resistance to extinction as a function of hours of food privation (preceding reinforcement constant) contains only four data points. The Mowrer-Jones curve of the number of extinction reaction as a function of work contains only three. Although the limits of "safe" extrapolation are more or less arbitrary, it is better to underestimate those limits than to overestimate them. This is particularly the case when dealing with the admittedly variable, and as yet quantitatively unexplored, data of psychology.

Nevertheless, it must be acknowledged that the desirable is not always the practicable. Provisional quantification—even when based on a restricted range of data—is preferable to no quantification at all. One can only admire Hull's scientific courage in sacrificing the security of vague qualitative generalization for the hazards of exact quantitative statement.

A fourth suggestion relates to an apparent implication of Hull's present analysis of inhibition. According to postulate 8, "the amount of net inhibition ($\dot{I}_R$) generated by a sequence of reaction evocations is a simple linear increasing function of the number of evocations" (300). Since the accumulation of inhibition is assumed to occur independently of whether the reaction is followed by reinforcement, the phenomenon of "inhibition of reinforcement" must be expected to occur. This means that empirical learning curves will rise only to the extent that the habit strength resulting from successive reinforcements exceeds the concurrent accumulation of inhibition. Recall now that habit strength is assumed to be generated according to a positive growth function and, thus, that the increments of habit strength progressively

decrease with each successive reinforcement. This consideration together with postulate 8 implies that when, at the limit of learning, increments of habit strength are becoming infinitesimally small, the concurrent increments of inhibition are becoming *progressively larger.*

The preceding deduction seems rather strange. It suggests that empirical learning curves carried well past the limit of practice should show a marked and positively accelerated fall as additional training is continued. Although the occurrence of inhibition of reinforcement effects has been fairly well verified, they do not seem to show either the magnitude or the form demanded by this deduction.

Another question arises in connection with one of Hull's symbols. The symbol "G" stands for: "a need reduction or a stimulus which has been closely associated with a need reduction; primary reinforcement; also a primary goal reaction" (404). This diversity of meanings is bewildering enough. The confusion is multiplied, however, when we find in a diagram of the chief variables of the system (Fig. 84, 383) that G is represented as one of the "observable antecedents" to which the intervening variables are anchored. Now, in so far as "G" stands for the *process* of need reduction or primary reinforcement, it must be regarded as itself an *intervening variable.* To the extent that "G" indicates a primary or secondary reinforcing *stimulus,* or an overt goal reaction, or any other condition which *gives rise* to need reduction, it can be considered an *independent variable* (antecedent condition). Thus, it would be helpful to analyze the rather flexible G into at least two components, one standing for the intervening variable, and the other for the parent independent variable. One possibility might be to let "C_G" stand for the antecedent conditions which give rise to need reduction (on the analogy to C_D). "G" might then stand for the intervening variable corresponding to the *process* of need reduction or reinforcement. The clarity and precision with which other symbolic distinctions of the system are worked out makes it all the more necessary to remedy this defect.

A final methodological consideration concerns Hull's views on the nature of intervening variables. Hull very reasonably assumes that legitimate intervening variables must be "securely anchored to directly observable events both as antecedents and consequences" (29). This is unquestionably true. Nevertheless, certain confusions are likely to arise if no attempt is made to examine the character of these "observable events" more closely. Intervening variables are rarely, if ever, *directly* linked to terms representing singular observations. Instead, they are typically related to independent and dependent variables which themselves are relatively high order scientific constructs. Characteristic independent variables in Hull's system are "stimulus" (S), "conditions

which give rise to drive" (C_D) or to need reduction (G), while among the dependent variables appear such constructs as "number of reactions preceding extinction" (n). These terms clearly cannot be equated with "directly observable events." Rather they represent empirical constructs which themselves must be defined in terms of their antecedent conditions and the consequent results. Ultimately, of course, such constructs must be reducible to classes of sentences containing "observable predicates." It should be emphasized that the independent and dependent variables appearing in most scientific laws have this character.

If Hull had made these considerations explicit, he could have dismissed the objections of extreme positivists to the use of intervening variables much more vigorously than he does at various places in the book. Hull's primary argument against positivistic procedure is that intervening variables represent a "convenience in thinking, i.e., an economy in the manipulation of symbols" (111). Of course they do. But the basic absurdity of the positivistic approach is that the constructs which are considered legitimate (independent and dependent variables) *have the same logical form* as the intervening variables which are held illegitimate. If intervening variables must go then, on the same grounds, most independent and dependent variables must go. This leaves all the brute observables one might desire, but no science.

Principles of Behavior arrives at a time when the withering *corpus* of the social sciences requires radical therapy. The war has seen a widespread resurgence of anti-scientific, metaphysical and theological speculation in the social sciences. There are some who mistakenly believe that the application of scientific methods to behavior and society has been tried and—witness the present social crisis—has failed miserably. Nothing could be further from the truth. If the social "sciences" have not realized their promise, this is primarily due to the *absence* of rigorous scientific method. The remedy is to make the social sciences more scientific, not less so. As a concrete step in this direction, the importance of *Principles of Behavior* cannot be overemphasized.

INTERACTION IN ANALYSIS OF VARIANCE INTERPRETED AS INTERCORRELATION

CHARLES C. PETERS
The Pennsylvania State College

In this article it will be shown that analysis of variance, reviewed by Garrett and Zubin as a research technique in the April, 1943, number of the *Psychological Bulletin* (1), can be readily pursued into a further analysis with much enrichment in its meaning, increased precision in its use, and a tie-in to the older statistical concepts. In particular, it will be shown that "interaction" is expressible in terms of the intercorrelation between classes, which reduces in the case of two classes to the long-familiar test of significance where matched groups are compared. But first it may be well to remark that "analysis of variance" as the term is customarily employed in tests of significance, is something of a misnomer which is likely to lead to misunderstanding. The variance of *a sample* can be analyzed into parts so long as one stays within the sample, as shown by Garrett and Zubin (1, pp. 234–235); but the technique as used in practice deals almost entirely with estimates of the *population* variance. That the population variance is not analyzed into constitutuent parts, which sum to a predetermined whole, is illusted by the fact that the "mean squares" (the estimates of the population variance) given on page 240 of the article do not add up to the 185.2 that would be estimated from the total sum of squares.* It is the *sum of squares* that is analyzed into parts; then from each constituent sum of squares that interests us a population variance is estimated by dividing by the appropriate n. Then these estimates are compared to ascertain whether they differ in magnitude more than chance fluctuation in sampling would justify. It would be much more enlightening to call the process *comparison* of variances. Otherwise when we bring this system into juxtaposition with other procedures where the term is employed in a different (and correct) sense, confusion is likely to result. (Compare, for example, 5, p. 33)

As a basis for seeing that "interaction" is expressible in terms of intercorrelation, let us run through the very simple derivation of the formula for average intercorrelation. Let $x_1, x_2, \cdots x_a$ be scores on a different measuring instruments (forms of a test, ratings by judges, etc.), each taken as a deviation from the mean of its series. Then, if σ_s^2 is the variance of the sums of the scores for the N individuals,

* It will be assumed that the reader of this article will get the Garrett and Zubin (1) article before him as he reads. Otherwise my article will be largely meaningless.

$$\sigma_s^2 = \frac{\sum(x_1 + x_2 + \cdots + x_a)^2}{N}$$

$$= \frac{1}{N}\sum(x_1^2 + x_2^2 + \cdots + x_a^2) + \frac{1}{N}\sum(x_1x_2 + x_1x_3 \cdots) .$$

The elements in the first parenthesis are variances of the several forms of the measures and those in the second parenthesis are correlations between the forms if each is multiplied by $\sigma_i\sigma_j$. If we assume equal variances, take the r's to be nearly enough alike to permit treating them as an average designated by r_{II}, and take cognizance of the number of each type that we would have in the complete expansion, we would have:

$$\sigma_s^2 = a\sigma_i^2 + a(a-1)r_{II}\sigma_i^2$$

$$(a^2 - a)r_{II}\sigma_i^2 = \sigma_s^2 - a\sigma_i^2$$

$$r_{II} = \frac{\sigma_s^2/\sigma_i^2 - a}{a^2 - a} . \tag{1}$$

We can put this in terms of the means of the scores instead of their sums by observing that, since a elements enter into each sum, $\sigma_s^2 = a^2\sigma_m^2$. So

$$r_{II} = \frac{a^2\sigma_m^2/\sigma_i^2 - a}{a^2 - a} = \frac{a\sigma_m^2/\sigma_i^2 - 1}{a - 1} . \tag{2}$$

This is the formula for the average intercorrelation among columns of a matrix table when m represents the means of the values on a row and σ_i^2 is the sample variance of a column assuming equal variances in the several columns. Since the assumption of equal variances is seldom perfectly fulfilled, we determine the column variance by taking a weighted average over all the columns, just as is the practice in analysis of variance; only, we stay with the *sample* variance.

Since no detailed table illustrating analysis of variance into three parts, with consequent interaction, is given by Garrett and Zubin, and since we imperatively need one to carry our point, let us set up the accompanying simple hypothetical one.

Suppose that five different methods of teaching are being tried out in each of six school systems. The mean made by each school is entered in its respective cell. For the present we shall pay no attention to any special characteristic of the rows but center attention only on the columns.

MEAN SCORES BY SCHOOLS AND BY METHODS

Schools	Methods						
	1	2	3	4	5	Sums	Means
1	25	27	24	28	22	126	25.2
2	24	32	29	26	24	135	27.0
3	31	35	27	36	26	155	31.0
4	40	45	33	42	30	190	39.0
5	43	50	38	46	33	210	42.0
6	45	48	40	52	36	221	44.2
Sums	208	237	191	230	171	1037	34.6

We shall be comparing the success of the five methods by comparing the means of the five columns. This we shall do by "analyzing the variance" into two parts, one associated with the means of the several columns and the other with the "error" element. To do this we compute a variance of the means (standard-deviation-squared) and estimate a population variance of the individual *schools* by multiplying this by the ratio $kn/(k-1)$, where k is 5 (the number of columns) and n is the number of entries in each column, which is 6 in our table. In practice we would proceed in a different but equivalent way: we would divide the sum of the squares of the means taken as deviations from the total mean by $(k-1)$, each mean having been taken n times. Then we would make another estimate of the population variance from the scatter within the several columns, taking each score as a deviation from its own column mean. In column 1 the sample variance would be

$$\sigma^2 = \frac{\sum(y-\bar{y}_c)^2}{n_c}$$

and, because we estimate the population variance by dividing the sum of squares by $N-1$ instead of by N (2, p. 70), the estimate of the population variance would be

$$s_{c_1}^2 = \frac{\sum (y-\bar{y}_c)^2}{n_c} \cdot \frac{n_c}{n_c-1} = \frac{\sum (y-\bar{y}_c)^2}{n_c-1}. \qquad (3)$$

But since we wish a more stable estimate than that from a single column, we take a weighted average over all the columns, then divide by the sum of the weights, and have:

$$s_c^2 = \frac{1}{\sum_1^k (n_c-1)} \sum_1^k (n_c-1) \cdot \frac{(y-\bar{y}_c)^2}{n_c-1} = \frac{\sum\sum (y-\bar{y}_c)^2}{N-k}. \qquad (4)^*$$

* There is a more mathematically rigid method of deriving this formula, by the use of calculus, which arrives, however, at the same result. See (2) pp. 349–350.

We now have two estimates of the population variance, which are known to be independent. If there is no law present which involves that the mean by one method of teaching is higher than by another, these estimates will be equal within the limits of sampling fluctuation. It would have been feasible to test this difference by the usual subtraction method with suitable tables of the distribution of the differences in standard terms. But R. A. Fisher chose to test the difference by *dividing* one by the other instead of subtracting, and tabled certain levels of probability of getting half the natural logarithm of this quotient. Snedecor translated these log values directly into the corresponding ratios between the variances, which he called F.

We spoke above of getting the "within the column" variance directly in the usual manner of computing a variance from the sum of squared deviations from the mean of an array. But while this is entirely feasible and is desirable as an arithmetic check, it is also feasible to get the necessary sum of squares by subtracting the sum of squares "between means" from the total sum of squares. This grows out of what happens when we set up the identity,

$$(y - \bar{y}) = (y - \bar{y}_c) + (\bar{y}_c - \bar{y})$$

square the polynomial, and sum over the N of the sample. The cross products vanish (sum to zero), and we have

$$\sum (y - \bar{y})^2 = n \sum (\bar{y}_c - \bar{y})^2 + \sum \sum (y - \bar{y}_c)^2 .$$

The two sums of squares that we drew upon for the estimate of the population variance as discussed above are the two on the right in the equation. It will be seen that these two partial sums of squares aggregate the one on the left. This illustrates the additive character of the sums of squares. It will also be seen that the denominators by which we divided our two sums of squares $(k-1)$ and $(N-k)$ add to the $N-1$ by which we divide the total sum of squares to make the optimum estimate from it. That illustrates the additive character of the "degrees of freedom."

The arithmetic values accruing from our table above when thus analyzed into two parts are as follows:

	sum of squares	*df*	"mean square" (Est. of Pop. V.)
Between methods	497	5−1	124.25
Within method columns (residual)	1740	30−5	69.6
Total	2237	30−1	—

Thus the estimate of the population variance from between the means is 124.25/69.6 = 1.78 times that from within the columns (the

"residual" or "error" factor). Reference to Snedecor's table (4, page 185) shows that, with 4 and 25 degrees of freedom, it takes an F of 2.60 to be significant at the 5 per cent point and one of 3.86 to be significant at the 1 per cent point. So no significant differences among the means of methods is suggested by this over-all comparison.

But now let us take cognizance of the arrangement by rows. Let us suppose that all of the schools of row 1 are from cities of a certain size of population; and let us suppose a similar meaning for each of the other rows. Let us further suppose that these rows represent cities of increasing size as we go from row to row, although such hierarchical arrangement is not necessary for our purpose. It is only necessary that the schools be matched on some criterion along the rows.

We now take cognizance of that matching by analyzing into three parts instead of two. To do this we compute a further sum of squares from the means of the *rows*, just as above we did so from the means of the columns. Since all the sums of squares must add up to the old total,* this new sum of squares must be taken out of the residual. But then the new residual sum of squares must be divided by $(k-1)(n-1)$ instead of by $(N-k)$ to make the optimum estimate of the population variance. The arithmetic values now are:

	sum of squares	*df*	Est. of Pop. V.
Between methods	497	$5-1$	124.25
Between schools	1588	$6-1$	317.6
Residual	152	$(5)(4)$	7.6
Total	2237	$30-1$	—

When now we make an F ratio out of the residual estimate of the population variance (7.6) and the between methods (124.25), we get an F of 16.3, which is very highly significant. If we had made a two-part analysis of schools without considering methods, we would have obtained an F of 12.2; so that the three-part analysis shows much higher significance than either alone. Since there is a joint effect that differs from either alone, it is said to be due to the interaction between the two: in this case, to the manner in which size of school affects the success of the methods.

We are now ready to show that this interaction is definitely expressible in terms of the average intercorrelation among the columns.

* This arises out of squaring and summing the identity,

$$(y-\bar{y})=(\bar{y}_c-\bar{y})+(\bar{y}_r-\bar{y})+[(y-\bar{y}_c)-(\bar{y}_r-\bar{y})]$$

which, since the cross products sum to zero, gives:

$$\sum(y-\bar{y})^2=n\sum(\bar{y}_c-\bar{y})^2+k\sum(\bar{y}_r-\bar{y})^2+\sum\sum[(y-\bar{y}_c)-(\bar{y}_r-\bar{y})]^2$$

Let us get before us again the analysis of sums of squares involved above. When analyzed into two parts

$$\sum (y - \bar{y})^2 = n \sum (\bar{y} - \bar{y}_c)^2 + \sum \sum (\bar{y}_c - y)^2 . \quad (5)$$

The analysis into three parts gives

$$\sum (y-\bar{y})^2 = n \sum (\bar{y}-\bar{y}_c)^2 + k \sum (\bar{y}-\bar{y}_r)^2 + \sum \sum (y-\bar{y}_c-\bar{y}_r+\bar{y})^2 . \quad (6)$$

Now (6) differs from (5) only in the fact that $k\sum(\bar{y}-\bar{y}_r)^2$ is taken out of the residual. The fact that this is $nk\sigma^2_{m_r}$ should at once suggest from (2) that it is expressible as a function of the intercorrelation. In these new terms the a of (2) is k, the σ_m^2 is $\sum(\bar{y}_r-\bar{y})^2/n$, and the $\sigma_t^2 = \sum\sum(y-\bar{y}_c)^2/nk$. So in our new terms, by some algebraic juggling from (2),

$$k \sum (\bar{y}_r - \bar{y})^2 = \frac{(k-1)\sum\sum (y - \bar{y}_c)^2 r_{II}}{k} + \frac{\sum\sum (y - \bar{y}_c)^2}{k} .$$

This, as said, is taken from the residual of (5) to give (6). There is left in the residual

$$\sum\sum (y - \bar{y}_c)^2 - \frac{(k-1) r_{II} \sum\sum (y - \bar{y}_c)^2}{k} - \frac{\sum\sum (y - \bar{y}_c)^2}{k} . \quad (7)$$

If the r_{II} is zero, this leaves

$$\frac{(k-1)\sum\sum (y - \bar{y}_c)^2}{k} .$$

To estimate the population variance this new residual must be divided by $(k-1)(n-1)$, which gives

$$s^2 = \frac{\sum\sum (y - \bar{y}_c)^2}{k(n-1)} = \frac{\sum\sum (y - \bar{y}_c)^2}{N-k} .$$

But this is exactly the same as the estimate from the residual in analysis into two parts. So the change from interaction is nothing if the average intercorrelation of the columns is zero.

If r_{II} equals 1.00, we have from (7)

$$\sum\sum (y - \bar{y}_c)^2 - \frac{(k-1)\sum\sum (y - \bar{y}_c)^2}{k} - \frac{\sum\sum (y - \bar{y}_c)^2}{k} = 0 .$$

Thus if the columns are perfectly intercorrelated, the estimate of the population variance from the residual is zero and the F is infinite. Whenever the columns are positively intercorrelated, the error (residual) estimate is descreased and the reliability of the determination

increased; when the average intercorrelation is negative, the residual estimate is increased and the reliability lessened; and when the average intercorrelation is zero, the estimate is unchanged.

By putting the elements of (7) over a common denominator and simplifying, we can place the residual directly in correlation form. We have for the new residual.

$$\frac{(k-1)(1-r_{ll})\sum\sum(y-\bar{y}_c)^2}{k} \tag{8}$$

The population variance is estimated from it by dividing by $(k-1)(n-1)$, giving

$$s_{rk}^2=\frac{(1-r_{ll})\sum\sum(y-\bar{y}_c)^2}{k(n-1)}=(1-r_{ll})\frac{\sum\sum(y-\bar{y}_c)^2}{N-k}=(1-r_{ll})S_c^2 \tag{9}$$

Thus the estimate of the population variance after taking account of interaction is $(1-r_{ll})$ times as great as it was before.

Often the data permit a further estimate of the population variance with which to compare the interaction: one from within the cells. Each of the 30 schools in our illustration had a number of pupils who individually got scores. If U is the number in a cell (a school) and the deviations are taken from the mean of the cell,

$$s^2=\frac{\sum(y_i-\bar{y}_i)^2}{U-1}$$

But the estimate is unstable if taken from a single cell, so we shall take an average over all the cells, weighting each for its n and dividing by the sum of the weights, as in (4). Thus we have an estimate of the population variance of individual observations (scores of pupils) by the formula

$$s^2=\frac{\sum\sum(y_i-\bar{y}_i)^2}{N(U-1)}$$

where N is the number of schools and NU the total population of pupils in the sample. This estimate of the population variance of individual pupil scores can be compared with the estimates of means of schools with which we dealt above by multiplying the latter values by U. It is from 8 such individuals in each of 8 cells that Garrett and Zubin (1, p. 246) get their $64-8=56$ degrees of freedom.

This showing that interaction is expressible in terms of intercorrelation would be merely pedantic if it were not for the fact that it can lead us on to a much enriched interpretation of what we have and can suggest more precise further studies. We can direct our attention to the

behavior of the columns in this respect and see, in the case of our illustration, that all the methods tend to be more successful as the cities in which they are employed are larger. Of, if the cells have been placed in rows on some other criterion, we can perceive that the scores from all the schools of a row tend to rise or fall together and, even if the rows have not been ordered in some hierarchical manner to begin with, we can speculate about what the characteristic is on the basis of which the methods are intercorrelated and can then pursue further the suggested hypotheses. A corresponding further analysis and pursuit would be possible if the average intercorrelation were negative, suggested by the increase in the residual from interaction. Or inspection of the columns might show that some pairs are positively intercorrelated and some negatively. To this situation the interaction variance would not do justice, because the negative and the postive r's would cancel each other. Of course, if the null hypothesis were fulfilled, columns would not thus behave individualistically; but the facts are not changed by making an assumption. Common sense would suggest that we override the over-all evidence from interaction, lift these cases out for more analytic treatment, and pursue by more precise methods the hypotheses suggested by the observation.

But the fact that interaction is expressible in terms of intercorrelation is especially illuminating when there are only two columns. When we analyze into three parts, the new error factor becomes, as examination of (9) will show,

$$s_e^2(1 - r_{12}) \tag{10}$$

where s_e^2 is the estimate of the population variance made from two columns treated as independent, estimated as follows:

$$s_e^2 = \frac{\sum (y_1 - \bar{y}_1)^2 + \sum (y_2 - \bar{y}_2)^2}{N_1 + N_2 - 2}$$

and r_{12} is the correlation coefficient between the two columns. If these two columns were being treated by the t technique, which is always a feasible alternative with two groups, the t would be for the independent groups which are implied where we analyze into only two parts, and the N's are equal as they must be if we are later to analyze into three parts,

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_e^2}{N} + \frac{s_e^2}{N}}}. \tag{11}$$

When now we analyze into three parts so as to take account of the interaction, our new estimate of the population variance for the denomina-

tor becomes that indicates in (10). Substituting this in (11) we have

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_c^2(1 - r_{12})}{N} + \frac{s_c^2(1 - r_{12})}{N}}}.$$

The two elements under the radical can be segregated into a shape that has a very familiar form, as follows:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_c^2}{N} + \frac{s_c^2}{N} - \frac{2r_{12}s_cs_c}{N}}}$$

$$= \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\sigma_{m_1}^2 + \sigma_{m_1}^2 - 2r_{12}\sigma_{m_1}\sigma_{m_2}}}.$$

So the interaction technique gives us our old formula for the standard error of the difference between matched groups where the correlation is (rightly) taken into account, and where a common variance is estimated from the two groups. But here there would have been available a more precise alternative involving even less labor: it is to *subtract* the paired variates instead of adding them, from which paired differences the standard error of the difference between the means is (2, p. 164–165)

$$\sigma_{m_1 - m_2} = \frac{\sigma_{y_1 - y_2}}{\sqrt{N - 1}}.$$

This is more precise because it makes fewer assumptions; it does not assume equal variabilities in the two arrays since the paired differences are normal variates having as such a variance with the usual sampling distribution without regard to the relative magnitudes of the variances of the two arrays which yielded them as differences. And the t from it has a known "exact" distribution; it was, in fact, the type of material for which "Student" originally developed his distribution.

The fact that interaction is expressible as intercorrelation ought to suggest that analysis of variance applied to testing the reliability of a test, cited by Garrett and Zubin on page 249, would yield to a further illuminating interpretation. And it does. The "between individuals" sum of squares is N times the variance of the averages of the paired items, and (2) shows that this has the makings of the coefficient of correlation between the tests. In fact a little algebraic juggling shows that the difference between the sum of the squares for "between individuals" and the sum of squares for interaction divided by the sum of these two

sums is the reliability coefficient of the test. But this assumes equal variance in the two arrays, a rough assumption. The F_2 test is a form of testing whether the obtained coefficient differs significantly from zero, seldom an important test for a reliability coefficient. The F_1 test is precisely the t test (actually t^2) of the significance of the difference between the means of the two forms where the denominator of the t ratio is the one that correctly takes account of the coefficient of correlation between the arrays but assumes a common variance for the two arrays. The t test taken from the difference between the two means divided by the standard deviation of the paired differences over the square root of $(N-1)$, as said in our preceding paragraph, avoids this assumption, is arithmetically easier, and has an "exact" distribution.

Garrett and Zubin (1, p. 250) say that the difference between the sum of squares "between individuals" and the sum of squares for interaction divided by the sum of these two sums (which they put in sample variance terms) is the *intraclass* correlation while, in our last preceding paragraph, we called it the *interclass* correlation. The intraclass correlation differs from the interclass only in the fact that the former takes all deviations from a common mean while the latter takes the deviations in each array from the mean of its own array (2, p. 201). It will be worthwhile to give the evidence why I believe the analysis of variance method gives the interclass correlation with assumption of equal variances rather than the intraclass.

For two arrays the interclass correlation, assuming equal variances is (2, p. 103):

$$r_{12} = \frac{2\sigma_m^2}{\sigma_t^2} - 1 = \frac{2\sum m^2/N}{\sum\sum (y - \bar{y}_c)^2/2N} - 1$$

$$= \frac{4\sum m^2}{\sum\sum (y - \bar{y}_c)^2} - 1 = \frac{4\sum m^2 - \sum\sum (y - \bar{y}_c)^2}{\sum\sum (y - \bar{y}_c)^2}.$$

But $\sum\sum(y-\bar{y}_c)^2$ is the sum of squares from within classes when we analyze into two parts. When we analyze into three parts and thus get an interaction sum of squares, the old residual sum is broken into a new residual sum and a new "between means" sum (here into the "interaction" sum and the "between individuals" sum), as shown in (5) and (6). Hence $\sum\sum(y-\bar{y}_c)^2$ can be replaced by these two parts, which for simplicity we shall designate $2\sum m^2$ and $\sum w^2$. So we have

$$r_{12} = \frac{4\sum m^2 - 2\sum m^2 - \sum w^2}{2\sum m^2 - \sum w^2} = \frac{2\sum m^2 - \sum w^2}{2\sum m^2 - \sum w^2}.$$

(In the language of analysis of variance, the sum of squares between

means is $k\sum m^2$, and k is 2 here in the case of two columns. Division of numerator and denominator by $2N$ yields the V's of Garrett and Zubin). Whether this is intraclass or interclass correlation depends upon whether the deviations in $\sum\sum(y-\bar{y}_e)^2$ are from a common mean or from the means of the respective columns. It is the universal practice in analysis of variance to take them in each column from the mean of that column; and that is why I believe this to be the old stand-by interclass correlation.

But surely we would not wish to introduce the intraclass correlation to describe the reliability of a test. For it would combine into a meaningless fusion the two very unlike features of difference between the means of the two forms and order of size of scores. We could not put the intraclass correlation to the practical uses to which we put the reliability coefficient, such as predict or regress with it or correct inter-function correlations for attenuation.

So far we have been talking about the interaction that arises by analyzing sums of squares into three parts when we have cells arranged in rows and columns that can be conceived of as making hierarchical arrangements, and we have shown that the interaction is then expressible in terms of intercorrelation. Thus there is brought together in each cell a different combination of the row and column factors. This same phenomenon could be pursued through analysis of variance into four parts in the Latin square and into five parts in the Greco-Latin square. Here there would be more multiplex combinations of factors. But the interaction could still be expressed in terms of correlations, since each new sum of squares of means of classes has the makings of a coefficient of correlation. But interaction effects are also studied in a more direct manner, as set forth by Garrett and Zubin on page 245. Here factors in various combinations are set up for direct experimental comparison. Each of these may be compared with a control, in which case there will be (k-1) comparisons, if k is the total number of cells. But also each may be compared with every other. Then the "degrees of freedom" will not add up to one less than the number of cells but to many more. But some of these comparisons will not be "independent"; for when both a and b have been compared with c, a and b have thereby a basis for comparison with each other. But Snedecor (4, p. 343) recommends, with his characteristic common sense, that the comparisons be made anyway if they are important to us.

The presentation of these results in tables is often baffling to the uniniated because of inadequate notation. One degree of freedom is assigned to each cell. But one degree of freedom in this connection always means that *two* means are being compared ($k-1=1$); it would be a great kindness to the reader if the writer would indicate clearly

which two, and not only the single one. And it will be wholesome for the reader to keep in mind that any test with one degree of freedom where variates are involved could, if one chose, be made by the *t* test for difference between means; for in the denominator of the *t* formula we want the best estimate we can get of the population variance, and to take this estimate from a combination of the many classes instead of only two does not disturb the fundamental nature of the *t* test. Even if the *F* test with two means is a little more convenient for penciling, and can be displayed more compactly in tables, thinking will be kept more clear if the reader will always keep firmly in mind that the thing that is going on is the long-familiar one of testing the difference between two means by comparing that difference to its standard error.

We have shown in this article that "interaction" is capable of further enriching interpretation beyond its surface meaning in terms of intercorrelation among the classes (columns or rows), and that such further analysis leads into the further testing of suggested hypotheses by more precise methods. We could make a similar showing with every one of the Fisher techniques explained in the Garrett and Zubin article.* They are *over-all* procedures, designed to discover quickly those hypotheses which may be dismissed without further labor, usually conducted as preliminary researches with small samples. Or, if they are done with adequately large samples, they are a first-over investigation for purposes of orientation to give leads for the more analytic investigation of suggested hypotheses with the same data. They have healthily encouraged systematic sampling, and replication in sampling, which have hitherto received too little attention in our field. They also lend themselves to compact presentation of data and to the mechanization of the arithmetic so that it can be readily assigned to clerical workers. But for this latter advantage we pay a very heavy price—too heavy—if it brings about the consequence that workers operate only with the mechanics while failing to understand clearly the fundamentally simple things they are doing, and if the presentation of the studies in publication takes a form that the bulk of the readers can not fully understand and pass critical judgment upon. In particular, the additive character of sums of squares and of degrees of freedom, which is merely an algebraic by-product rather than a fundamental principle, is likely to invite procedures that rest on magic and authority. It may be a little more awkward in exposition and in arithmetic to get one's eye every time on the sum of squares for a class of observations or of statistics and compute

* I have shown elsewhere that analysis of variance is a form of the unbiased correlation ratio (2, p. 353–357) and that analysis of covariance can be simply and profitably set up as a regression technique for matching groups for experiment without loss of population (3).

an estimate of its population variance by dividing by $(n-1)$ or some function of that, but that procedure keeps the process clearly understandable. When we do hundreds of these each day, it will be time to introduce the arithmetic short-cuts and the other advantages of mechanistic procedure.

And when, instead of stopping with the over-all tests of significance as if they were the end of the story rather than its beginning, we continue into the more analytic procedure that lies beyond, and to the more precise pursuit of the suggested hypotheses, we shall find ourselves able to tie in what we have with the long-growing central stem of classical statistics. But when we go beyond the Fisher statistics we shall not merely return to the old classical statistics; we shall advance into some further developments and refinements of precision that lie as definitely beyond the present levels of the "new" statistics as it now lies beyond the "old." And, strange though my promise may seem, this further development can be achieved within the understanding of the rank and file of readers who have hitherto understood statistics, because they will synthesize on a higher level and generalize the "old" and the "new" instead of leaving the present gap between them.

BIBLIOGRAPHY

1. GARRETT, H. E., AND ZUBIN, J., The Analysis of Variance in Psychological Research, *Psychol. Bull.*, 1943, **40**, 233–267.
2. PETERS, C. C., AND VANVOORHIS, W. R., Statistical Procedures and Their Mathematical Bases, McGraw-Hill Book Co., 1940.
3. PETERS, C. C., A Method of Matching Groups for Experiment With No Loss of Population, *J. educ. Res.*, 1941, **34**, 606–612.
4. SNEDECOR, G. W., Statistical Methods, Collegiate Press, 1938.
5. WITTENBORN, J. R., Factorial Equations for Tests of Attention, *Psychometrika*, 1943, **8**, 19–35.

NOTE: SHALL AFFILIATED ORGANIZATIONS HAVE DIRECT REPRESENTATION ON THE PROPOSED COUNCIL OF REPRESENTATIVES?

JOHN GRAY PEATMAN
City College of New York

By a unaminous vote, the New York State Association for Applied Psychology, at its Annual Meeting on February 19, approved a Resolution requesting the Joint Constitutional Committee to eliminate the Board of Affiliates and nine Regional Representatives from the proposed APA reorganization, and to substitute therefore *direct representation* on the proposed Council of Representatives for organizations of psychologists now affiliated with the APA or AAAP.

The afternoon session considered the proposals of the Intersocietal Constitutional Convention for a reconstituted APA and dissolution of the AAAP. Dr. A. T. Poffenberger presided. Dr. Alice I. Bryan of the Joint Constitutional Committee outlined the present plan and discussed its implications. Miss Margaretta Weber and Dr. Herbert B. Williams presented a series of questions with respect to the effect of the plan on affiliated organizations of Applied Psychology.

In opening the session, Dr. Poffenberger emphasized the importance of the reorganization of our scientific and professional societies. He indicated that what we have at present is a result of 50 years of haphazard growth, and that it is the hope of those concerned with the problem that some kind of national pattern can be set up which will better serve the professional and scientific needs and standards of psychologists. He further emphasized that the present status of the reorganization plan was in the nature of a series of proposals, and that final action is yet to be taken by the membership of the APA and the AAAP.

Dr. Bryan pointed out that the impetus for the proposed merger arose out of the recognized need for establishing a permanent central office with a full-time Executive Secretary to function for the profession of psychology on a year-round basis. She further emphasized that one national association would not only eliminate over-lapping memberships and multiplicity of dues, but would also enable the professional and scientific functions of psychologists to be carried on more expertly than at present. Dr. Bryan concluded her presentation by pointing out alternatives to the present proposals; namely, the maintenance of the *status quo*, the postponement of a plan of unification until after the war, or the modification of present proposals between now and the final voting.

Miss Weber outlined in some detail the advantages and disadvantages of the present proposals. She indicated that there would doubtless be many advantages to be derived from a strong centralized national organization of psychologists, so long as the professional needs and interests of regional and State groups were not neglected. In preparation for the meeting Miss Weber pointed out that she had discussed the proposals with various members of the Association. Some members, with whom Miss Weber had talked, wondered whether the

present plans were sufficiently mature to be ready for acceptance or rejection by this year. Others questioned the fairness of a vote on the proposition this year with so many psychologists serving in the Armed Forces. Some members were concerned lest the proposed national organization be controlled by the so-called academic psychologists. Miss Weber also indicated that the status of Associate members in the merged organization and the prerogatives of the proposed Committee on Committees had not been made sufficiently clear in the tentative By-Laws published in the November 1943 issue of the *Psychological Bulletin*. In concluding, Miss Weber questioned the soundness of the proposed Regional Representatives to the Council of Representatives, in view of the fact that such delegates would not actually represent functioning organizations of psychologists.

Dr. Williams pointed out that the proposed merger of the APA and the AAAP no doubt could be accurately described as a proposal for a reconstitution of the present structure of the APA. A fundamental question, he indicated, is whether such a reconstitution of the APA would advance Psychology as a profession. He doubted whether a new APA would be as successful in this regard as the present AAAP and existing regional organizations in the applied field. Dr. Williams felt that the present proposals might lead to a national organization which would function for psychology mainly as the American Association for the Advancement of Science functions for many areas, rather than result in a national association devoting its energies to the development of Applied Psychology. Pointing to the accomplishments of the American Association of Social Workers, Dr. Williams emphasized that this organization of professional people had, in a relatively short time, established a growing institution in the non-academic, applied field. Some states have already enacted legislation which require that a social worker be licensed in order that definite standards of competency be maintained by those practicing in this field. Perhaps a reconstituted APA might be able to achieve some such recognition for psychologists as professional people. If so, he added, he would like very much to see a full consideration of this possibility before final action is taken on the proposed reorganization. In concluding, Dr. Williams indicated that the war has revealed a scarcity of psychologists competent to meet many problems which properly lie within the field of Applied Psychology. Much of the work now being done by social workers in connection with personnel selection, personality studies, guidance, etc., would more properly belong to Applied Psychology. No doubt such work would be carried on by more psychologists had we the professional schools in Applied Psychology. Unfortunately, most of our graduate schools have not developed professional training programs to meet such needs. If something is not done about this situation, either by academic psychologists or by professional psychologists in the applied field, psychology will continue to "miss the boat." This will be true not only with regard to the various areas mentioned but also in connection with the Rehabilitation Program now being organized to serve the needs of men and women returning from the Armed Forces.

During the open discussion following Dr. Williams' remarks, several members voiced the opinion that if the plan for a new APA is adopted, and if such a national organization is really to represent psychologists in all parts of the country and serve their professional needs, then at least the already established regional associations in the applied and non-applied fields should have *direct representation* on the congress of the national organization. Otherwise, a highly centralized national organization might easily lose sight of the State legislative and other special problems of local constituents.

Adoption of Resolution at the Business Meeting

At the business meeting of the Association, Dr. Peatman opened the discussion by presenting a resolution which would modify the proposed By-Laws by permitting regional organizations now affiliated with the APA or AAAP to have direct representation on the proposed Council of Representatives. He indicated that, in view of the afternoon discussion and the tentative status of the proposed By-Laws, the Association was hardly in a position to go on record as either approving or disapproving the proposed reconstitution of the APA and attendant dissolution of the AAAP.

In support of the resolution, it was pointed out that direct representation on the congress of a national organization would serve to implement the development of the professional needs and interests of the 17 regional associations now affiliated with the APA or AAAP. Furthermore, direct representation would give a fair amount of really organized regional balance to an otherwise highly centralized national organization. The resolution would substitute direct representation of affiliate organizations for the proposed nine Regional Representatives and a Board of Affiliates. The so-called Regional Representatives would hardly serve any satisfactory purpose, inasmuch as they would be responsible to no active regional associations. Furthermore, the interjection of a Board of Affiliates between affiliated organizations and the congress of a national organization would not only be an indirect method of representation but would be cumbersome and inadequate to serve the needs of affiliate groups.

During the discussion it was further pointed out that the adoption of the resolution would serve to clarify the general structure of the proposed organization. There would be representation in the congress not only of the various fields of Psychology (by the Division Representatives), but there would also be minority representation (by Affiliate Representatives) for the special professional needs of geographically localized groups. It was emphasized that, in the long run, the affiliate organizations would in all likelihood be the particular groups of professional psychologists who could more frequently gather, rub elbows and exchange ideas. The direct representation of affiliate organizations on the Council of Representatives should therefore prove to be a continuous vitalizing influence for a national association, and at the same time facilitate the recognition by State legislatures and by local institutions of the work and standards of both local and national organizations. It was pointed further out that the intent of the Resolution is to permit regional associations with 50 or more members (the same minimum proposed for Interest Divisions), and now affiliated with the APA or AAAP to become Charter Affiliates of the reconstituted national association at the time it is established. Should all of the existing affiliate organizations have the minimum number of members deemed appropriate and wish such affiliation and consequent direct representation on the Council, there would be an initial maximum of 17 Affiliate Representatives. Even if the Council should prove to have an initial minimum of but 60 Division Representatives, an initial possible maximum of 17 Affiliate Representatives would not only be a minority of the Council as a whole but would comprise only five or six (at most) more representatives than provided for under the plan of nine Regional Representatives and a minimum of two representatives from the Board of Affiliates.

Whether new regional groups of psychologists could affiliate with the national association after its reconstitution would, of course, depend on a vote of the entire Council. Consequently there would be no danger of Affiliate Representatives running away with or swamping the Council membership. It was pointed out that the machinery for the admission of new affiliate organizations or for the discontinuance of old ones could be similar to the By-Law structure already proposed for the admission and discontinuance of Interest Divisions.

The additional suggestion was made, during the discussion of the resolution, that a provision be added recommending non-pluralistic voting for Affiliate Representatives. It was also emphasized that the present and future functions of the Board of Affiliates of the AAAP could well be taken care of by a standing committee of the National organization. The membership of such a committee could be drawn from the Affiliate Representatives on the Council; in fact, the special professional interests of affiliate organizations in applied psychology could be served by a standing committee on Affiliate Organizations of Applied Psychology.

The following resolution was then put to a vote and unanimously adopted:

WHEREAS, The proposed By-Laws appropriate to a reconstituted American Psychological Association include a Board of Affiliates to represent affiliated organizations and also include nine Regional Representatives,—a structure of organization that would give rise to needless duplications of effort and to inefficiency in administration; and

WHEREAS, Active organizations of psychologists, bound together by interests and professional work within geographical divisions, such as several states, a state, or a metropolitan area, would under the proposed By-Laws have no direct representation on the proposed Council of Representatives; and

WHEREAS, The proposed Interest Divisions of a Reconstituted American Psychological Association would be organized on national lines and on the basis of more restricted functional activities and interests in psychology of a minimum of fifty (50) Associates and Fellows of the Association,

THEREFORE, BE IT RESOLVED, that the New York State Association for Applied Psychology requests the Joint Constitutional Committee to eliminate the proposed Board of Affiliates and the nine Regional Representatives and substitute therefor, on a permissive basis, Affiliate Representatives to the Council of Representatives for organizations of professional psychologists within geographical divisions, such as several states, a state, or a metropolitan area, and composed of a minimum of fifty (50) members, each such affiliate organization to be represented in the proposed Council of Representatives by one Representative, who shall be an Associate or Fellow of the national Association and who shall be chosen from and elected by the members of such affiliate organization, provided, however, that persons who are members of more than one such affiliated organization shall have the privilege of voting for the candidates to the Council of Representatives of but one affiliated organization at the time of nomination and election of Affiliate Representatives.

NOTE ON PSYCHOLOGY IN OCCUPIED SHANGHAI

C. HART WESTBROOK*

Early in the Sino-Japanese War, the Japanese took possession of the grounds and buildings of the University of Shanghai and held them for more than two years before returning them to their owners, the Northern and the Southern Baptist Foreign Mission Board of America, who were neutrals at that time. But the apparatus in the psychological laboratory, as with most of that in the physics, chemistry, and biology laboratories, was looted or destroyed. During the year 1937–38 the University, therefore, was forced to conduct its classes in elementary General Psychology without any laboratory work.

In the Fall of 1938, the St. Johns University of Shanghai, the Soochow University, and the Hangchow Christian College, began to cooperate with the University of Shanghai in the conduct of certain courses, particularly in those which involved the use of laboratories and libraries. At the same time the Committee for Christian Colleges and Universities in China, from its headquarters in New York City, came to the assistance of the universities in the purchase of such apparatus for our common use as could be obtained abroad or could be constructed under our supervision in Shanghai. I served as the Chairman of the Laboratory Committee for Psychology. Out of the lump sum available some funds were allocated for psychological apparatus. By hunting out little shops which were hidden away here and there in Shanghai, I was able to secure the construction of a number of simple pieces for the laboratories in General and Educational Psychology; we also purchased a few pieces from abroad. All four institutions had access to this apparatus and the laboratory in which it was placed. To the best of my knowledge, this equipment is still available for use by St. Johns University and the Hu Kiang Institute. The latter is a temporary organization set up to function after the University of Shanghai's Board of Directors suspended the University in February,

* With the author's permission, this note is taken from a letter written to the Editor of the *Psychological Bulletin*, who felt that the contents were of general interest. Professor Westbrook says that after the occupation he knew less about what was happening in "free China," than we in America can learn from the Shanghai *Evening Post* and *Mercury*, which are now published weekly in New York City. Hence, he speaks only for psychology in a limited area of China under conditions of occupation. Dr. Westbrook, who received his Ph.D. in Educational Psychology from Harvard in 1920, was formerly Professor of Psychology and Education in the University of Shanghai. For seven months he was interned in the Japanese internment camp at Pootung. He returned to the United States on the *Gripsholm* on December 1, 1943.

1942. The two other cooperating institutions disbanded about the same time and moved to free China. I had about eighty students each year in the laboratory courses at the University of Shanghai, Hangchow College had about sixty, and St. Johns had perhaps forty or fifty. I turned over the laboratory to Professor Chang, formerly of the Great China University, who took up my work in the Hu Kiang Institute, when I had to drop out before being interned by the Japanese at Pootung, across the Hwangpoo River, Shanghai.

Separately, the four Christian institutions conducted their own courses in General and in Educational Psychology, as well as in Child and Adolescent Psychology. But in order to save the time of the instructors and to broaden the list of offerings, they cooperated in courses in Experimental, Abnormal, and Social Psychology, and in Mental Hygiene. Professor J. H. Wiley (for three years) and Professor Arthur Chen, and Mr. C. T. Djang, both of the University of Shanghai, gave courses in Social Psychology. Professor Florence Sherriff and Professor James H. Pott, of St. Johns, taught courses in Social Psychology and Abnormal Psychology, respectively. Acting President W. B. Nance of Soochow University, and Dean T'an of Hangchow College, gave courses in their own institutions in General Psychology, and in General and Educational Psychology, respectively. I supervised senior theses in different phases of psychology related to my courses for the students of the University of Shanghai. Apart from this research by seniors and except for some experimentation in methods of teaching, neither the staff members nor the students had the time or the equipment to conduct any psychological research above the requirements for the Bachelor's degree at our University. A similar situation, so far as I know, held in each of the three other cooperating institutions.

The private institutions organized and operated by the Chinese apart from the Christian institutions, were bombed early in the war with Japan (1937). Except for small "token" groups of teachers and students, they moved out into free China before conditions made further moving impossible. Hence, at present there is nothing to be reported from the groups left in Shanghai. Doubtless the Chinese institutions, in cooperation with the University of Nanking, the Tsing Hwa University of Peking, and other universities, have carried on some psychological work in their new locations in the West. The China National Research Institute, formerly of Nanking, which also moved out, must have carried its apparatus and library into the interior before the fall of Nanking and is probably continuing to do some solid work. At least we hope so.

American institutions and scholars will have a great opportunity to assist both the numerous privately owned Christian colleges and the

Government colleges and universities in rebuilding their laboratories and their libraries after the close of World War II. I should certainly like to make such a request for our own institution, as well as for all others in the East China area which have suffered great losses. Of particular importance will be files of scientific periodicals. I would urge psychologists to preserve their journals and to make them available after the war to the bombed out and damaged universities and libraries of the world, in which the need for basic scientific literature will be great.

PSYCHOLOGY AND THE WAR

Edited by

DONALD G. MARQUIS

CONTENTS

RESEARCH PROGRAM ON PSYCHOMOTOR TESTS IN THE ARMY AIR FORCES by *Staffs PRU No. 2* and Department of Psychology, School of Aviation Medicine 307

A BLINKER ADAPTATION OF THURSTONE'S 1943 CODE APTITUDE TEST by *H. B. Carlson* and *J. Rich* 322

RESEARCH PROGRAM ON PSYCHOMOTOR TESTS IN THE ARMY AIR FORCES

STAFF, PSYCHOLOGICAL RESEARCH UNIT NO. 2
*San Antonio Aviation Cadet Center, San Antonio, Texas**

STAFF, DEPARTMENT OF PSYCHOLOGY, RESEARCH SECTION
*School of Aviation Medicine, Randolph Field, Texas**

This article is the fourth in a series describing the Aviation Psychology Program of the Army Air Forces. The first article (4) presents a general outline of this program. The second and third articles (5, 6) deal with research activities at Psychological Research Units Nos. 1 and 3, Nashville, Tennessee, and Santa Ana, California. The present report treats the research activities of Psychological Research Unit (PRU) No. 2, San Antonio Aviation Cadet Center, San Antonia, Texas, and of the Department of Psychology, Research Section, School of Aviation Medicine, Randolph Field, Texas.

I. ORGANIZATION AND PERSONNEL

In the development and evaluation of psychomotor tests in the Aviation Psychology Program, the Department of Psychology and PRU No. 2 shared responsibility. Frequent meetings of the officers have been

* This article is based on the annual reports for the fiscal year of 1943, prepared by the staffs of Psychological Research Unit No. 2, and the Department of Psychology and covers the period from the beginning of the units to 30 June 1943. The materials were abstracted and edited in the Psychological Branch, Research Division, Office of the Air Surgeon, Headquarters Army Air Forces, Washington, D. C.

held for the purpose of mutual presentation, discussion, and decision regarding such problems as the design and evaluation of new psychomotor tests, details regarding the administration of research psychomotor tests, statistical procedures to be employed in the analysis of research data, and preparation of final research reports. The design and production of psychomotor tests is the primary responsibility of the Department. The administration of experimental and classification tests to cadets and the refinement of administration procedures are primary responsibilities of PRU No. 2. Both units undertook statistical analyses. These procedures apply to research projects undertaken jointly by the two organizations; both the Department and PRU No. 2 carry on independent as well as joint research projects.

In addition, other psychological units have contributed ideas and developed several psychomotor devices.

A. *Psychological Research Unit No. 2*

The principal functions and objectives of PRU No. 2 are as follows:

(1) the testing of aviation cadets and students and recommending them as pilots, bombardiers, or navigators on the basis of psychological aptitude tests, ratings of interests, and flying experience; and

(2) the collecting of data and carrying out of research relating to the construction, improvement, and development of psychological aptitude tests for this purpose.

In the fulfillment of the second of these functions PRU No. 2 has worked in close collaboration with the Department of Psychology.

The procedures for processing cadets at PRU No. 2 are the same as those described in the previous articles (**5, 6**). PRU No. 2 is under the technical supervision of the Air Surgeon, through the Surgeon, AAF Training Command. The beginning of the unit dates from the arrival of Major (now Lt. Colonel) Robert T. Rock, Jr. in November, 1941, who served as Director of the Unit until he became Director of the AAF Classification Center, San Antonio Aviation Cadet Center, on 25 March 1943. Major A. C. Tucker acted as Director during the last part of the fiscal year 1943. As of 30 June 1943, the personnel of the unit consisted of seventeen officers and ninety-two enlisted men. All of the officers were graduates of accredited colleges or universities with either an undergraduate major in psychology or graduate work in this field. A large proportion held advanced degrees in psychology. A majority of the enlisted men had the bachelor's degree with a major in the field of psychology, and approximately 20 per cent held advanced degrees. Rosters of the names of the officers and enlisted men who have been on duty in the Aviation Psychology Program at PRU No. 2 prior to 30 June 1943, follow.

Officers Who Have Been Assigned to Psychological Research Unit No. 2

Crannell, Clarke W.
Crawford, Meredith P.
Dailey, John T.
DuBois, Philip H.
Erickson, Stanford C.
Finch, Glen
Ghiselli, Edwin E.
Grether, Walter F.
Henneman, Richard H.
Hildreth, Glenn W.
Killian, Frank, Jr.
Kleinsasser, Alvin J.
Lepley, William M.
McClelland, William A.
Payne, Robert B.
Rock, Robert T.
Sperling, Philip I.
Tucker, Anthony C.
Vallance, Theodore R.
Wallace, S. Rains
Weitz, Joseph
Weston, Julien V.
Youtz, Richard P.
Zielonka, William A.

Enlisted Men Who Have Been Assigned to Psychological Research Unit No. 2

Adams, Jack D.
Ahner, Charles W.
Anderson, Carl L.
Anderson, Evans L.
Amundsen, Earl H.
Applezweig, Mortimer H.
Armitage, Stewart G.
Atkinson, Norman
Bacon, Franklin, Jr.
Baker, John A.
Bardsley, Roger E.
Barnette, Warren L., Jr.
Benson, Arthur L.
Berger, Emanuel M.
Berkowitz, Melvin
Bernard, Jack
Bjork, Alton J.
Boory, John F.
Bryan, Robert C.
Buck, Raymond E.
Burack, Benjamin
Chaplin, James P.
Charley, William M.
Chiozza, Joseph P.
Clark, Jerry H.
Conklin, Edmund H.
Corbin, Horace H.
Dahlin, Bernard C.
DeMott, John J., Jr.
Dodds, William H., Jr.
Dreikurs, Eric
DuBois, Donald B.
Duncan, Robert L.
Farber, Isadore E.
Farberow, Norman L.
Feiner, Arthur H.
Fenger, Fred T.
Finney, Ben C.
Fitzgerald, Don Casot
Fox, Richard F.
Froehlich, Clifford P.
Futransky, David L.
Galt, William E.
Geimer, Vincent D.
Gellman, Martin S.
Gibbs, William T.
Gillman, Robert D.
Glaser, Robert
Goodman, Ervin
Goodman, John K.
Goodrich, Joseph W.
Gordon, Philip L.
Gough, Harrison G.
Gresham, William W.
Grice, George R.
Grings, William W.
Haire, Mason J.
Hamborsky, Dwight K.
Hamilton, Charles E.
Harless, Byron E.
Harris, Robert E.
Harris, William E.
Heartsill, Walter C.
Heyns, Roger W.
Hodben, Frank E.
Hudson, George R.
Hufstader, William H.
Irion, Arthur L.
Jensen, Elmo A.
Jernigan, Austin J.
Jerome, Eldon K.
Johnson, George O.
Kandel, Alan D.
Kehoe, John A.
Keil, Jack Payne
Keller, Robert J.
Kirk, Marquis A.
Klapper, Joseph T.
Klebanoff, Seymour G.
Knight, Norton B.
Krugman, Herbert E.
Lane, Robert C.
Lanier, Julien A.
Lassman, Frank M.
Laurito, Angelo F.
Lawrence, Douglas H.
Lehner, George F. J.
Lerner, Joseph S.
Levin, Morton
Levine, Abraham S.
Levine, Robert
Lieberthal, Jerome I.
Linn, Leslie W.
Lubin, Ardie
Mahl, George F.
Matheny, William G.
McReynolds, Paul W.
Meissner, James H.
Menozzi, John
Miller, Robert B.
Mogin, Bert
Mollenkopf, William G.
Muller, Henry J.
Munson, Charles B.
Murphy, Robert E.
Niehaus, Stanley W.
North, Robert D.
Olsen, Raymond

Pasco, Robert C.
Patterson, Cecil H.
Payne, Riner C.
Payne, Robert B.
Philips, Charles L.
Phillips, Ewing L.
Poinc, Arthur E.
Pollard, Richard F.
Pomeroy, Donald S.
Preston, Harley O.
Prinz, Gerald R.
Rahm, Hayden H.
Ramsey, Glenn V.
Redlener, Joe
Robbins, Irving
Rochlin, Isaiah
Rohles, Fred H.
Roney, James G., Jr.
Rosenwald, Alan K.
Royce, Joseph R.
Schapiro, Jack M.
Schmierer, Hyman
Schoenrick, Carlos O.
Schrader, Everett N.
Schwartz, Stanley
Scrivner, Robert L.
Settlage, Paul H.
Shea, Daniel J.
Shimberg, Benjamin
Shlien, John M.
Sincoff, Philip Lee
Smith, Moncrieff H., Jr.
Spencer, Charles M.
Sprow, Allen J.
Standish, William A.
Steinert, John C.
Stevens, Evan R., Jr.
Sullivan, Cornelius D.
Thomas, Francis H.
Thompson, Charles E.
Traeger, Carl
Valin, Sigmund
Webb, Wilse B.
Weinberg, Edward H.
Weiss, Aaron
Whitesell, Jack C.
Wiener, Daniel N.
Wiley, John H.
Willerman, Benjamin
Williams, Malcolm J.
Williams, Meyer
Willmann, John M.
Wilson, Everett E.
Wisely, Harold M.
Woodruff, Joseph L.

B. *Department of Psychology*

The functions of the Department of Psychology may be described as follows:

(1) Research leading to the origination, construction and evaluation of new psychomotor tests for the selection and classification of aviation cadets.

(2) The design, redesign, and production of psychomotor tests that have been approved for use in the Aviation Psychology Program.

(3) Cooperative research with other departments of the Research Section, School of Aviation Medicine, especially in those research problems that involve the use of psychomotor tests.

The design and production of psychomotor tests for use in the classification of aviation cadets in the AAF Training Command has been the responsibility of the Department of Psychology since March 1942. Although not strictly a research function, this activity has occupied the personnel of the Department continuously and on occasion, exclusively. Tests that had been validated in research forms had to be modified for use in a mass testing program, special cycling devices had to be developed for the automatic control of testing intervals and other conditions in order that four aviation cadets could be tested simultaneously by a single examiner, and tests and control devices had to be designed to minimize the differences between different copies of the same test. The Department of Psychology performed these functions and in addition had responsibility for the procurement of the necessary materials and the direct supervision of the production of the tests for use in the AAF Classification Centers. By 1 November 1943 twenty-five "lines" of psychomotor tests, each "line" being composed of four copies of each of

six apparatus tests, had been delivered to the AAF Classification Centers, and Basic Training Centers.

No actual testing of aviation cadets has been performed in the Aviation Psychology Program at the School of Aviation Medicine. The School of Aviation Medicine, of which the Department of Psychology is a unit, is under the direct supervision of the Air Surgeon. Psychological research is coordinated through the Research Division of the latter office.

Major (now Lt. Colonel) Arthur W. Melton was commissioned as an Aviation Psychologist and reported for duty at the School of Aviation Medicine in January 1942 and has served as Chief, Department of Psychology, since that time. Three additional officers—Captain Judson S. Brown, Captain Jack Buel, and Lt. Robert M. Gagne—and one Senior Psychologist (Dr. R. B. Loucks) reported during 1942. All of the officers and Dr. Loucks have the Ph.D. in psychology and have had extensive research experience in the development and use of psychological apparatus. In addition, two officers of the Aviation Psychology Program, Captain Roger W. Russell and Lt. Wilbur S. Gregory, served as civilians for a short time in the Department.

II. Rationale of Research

The basic assumption on which the Department of Psychology and PRU No. 2 worked in the development of psychomotor tests was that success in pilot and bombardier training involves to an important degree certain perceptual-motor skills, and that the abilities required for the rapid learning of these skills vary among individuals and can be measured best by psychomotor apparatus tests. When the Air Surgeon was given the mission of establishing a classification testing program in February 1942, it was decided to include several psychomotor tests in the battery for the selection and classification of pilots, bombardiers, and navigators. No preparation had been made to meet the demand for a large number of such apparatus tests and the activities of the Department of Psychology and PRU No. 2 were chiefly concerned during the first eight months of 1942 with the particular problems of a mass psychomotor testing program.

The research program of 1942 was particularly concerned with the immediate adaptation of promising psychomotor tests for use in the mass testing program. To this end a number of tests that had been shown by other research agencies to have some validity for the selection of pilots were purchased or constructed and modified, the modifications being validated at PRU No. 2 against the criterion of success or failure in pilot, bombardier, and navigator training schools. Throughout this period the concern was with conditions of testing such as reliability of

scores, instructions, practice and demonstration procedures, time limits, design of equipment to automatize the administration of the tests, scoring methods and apparatus, sources of apparatus variance and examiner variance, optimal size of group to be tested simultaneously by a single examiner, and target sizes. During this period there was also an opportunity to administer tests of a number of different types to large groups of aviation cadets for validation purposes. The majority of such tests were purchased or borrowed from laboratories of psychology, although some were constructed at the School of Aviation Medicine.

After the standardization of psychomotor tests for classification purposes had been accomplished in the Fall of 1942, attention was directed in more concentrated form to the development of new psychomotor tests. It was believed, on the basis of evidence on hand at that time, that the most successful tests would be those which involved relatively complex visual-motor coordination and at the same time involved a heavy weighting of the learning factor. Although it is agreed that too much stress should not be placed on making psychomotor tests duplicate some actual situation or operation in piloting an airplane or sighting a bombsight, it was believed that the most successful tests would be those that were simplified samples of such situations or operations. The majority of the research tests procured or produced and tested for validation purposes during 1943 were tests of visual-motor pursuit functions. Visual-motor pursuit tests had shown promising validity and it was believed that visual-motor pursuit functions should be thoroughly explored as a guide for the development of pursuit tests for the classification battery. In this exploration an attempt was made to obtain or produce pursuit tests that involved both "compensatory" pursuit and "following" pursuit, simultaneous coordination with reference to the movements of from one to four visual objects, simultaneous coordination with reference to a single visual object that would move in one, two or three dimensions, "all or none" vs. graded scoring systems for errors in performance, and various amounts of verisimilitude to the actual task of piloting an airplane.

The next section of this article is devoted to a description of the major tests that have been developed or modified and subjected to validation testing and a later section of this article is devoted to a description of the special techniques for mass psychomotor testing that were developed.

III. Psychomotor Classification Tests

It should be noted in the following discussions dealing with psychomotor apparatus tests that have been tried out experimentally against the criterion of success or failure in training that a number of these devices are adaptations of tests that were extant prior to the Aviation

Psychology Program. Also, several devices recently originated by psychologists have been borrowed or purchased for validation trial at PRU No. 2 and for redesign by the Department of Psychology to adapt them to the requirements of a mass testing program. In general, the policy has been to give experimental trial to any psychomotor test, regardless of its origin, if it seemed to have likelihood of measuring some ability required for aircrew success.

A. *Psychomotor Classification Tests in Use 1 July 1943*

1. *S.A.M. Complex Coordination Test* (*Mashburn*). This test was originally developed in 1934 by Colonel Neely C. Mashburn at the School of Aviation Medicine and has been known as the Mashburn Automatic Serial Action Apparatus (2) and as the Constable Automatic Serial Action Apparatus (7). This test was shown to be valid for the selection of pilots in a report presented by Major C. R. Glenn (now Brigadier General Glenn, Surgeon of the AAF Training Command) in 1935 (1) and in more recent studies accomplished under the auspices of the National Research Council Committee on the Selection and Training of Aircraft Pilots. This test has shown high validity for the selection of pilots in numerous large populations tested in the Army Air Forces during the last two years.

The S.A.M. Complex Coordination Test now in use in the Aviation Psychology Program of the Army Air Forces (Figure 1) is essentially the Mashburn Automatic Serial Action Apparatus with improvements designed to increase the dependability of individual pieces and the uniformity of testing conditions with different copies of the test. The candidate sits in front of three double rows of lights and by manipulating a stick and a rudder causes the apparent position of a green light in each of the three rows to match the apparent position of a red light in each row. When the three green lights have been made to match the three red lights simultaneously for a period of one-half second, a new pattern of red lights appears. The candidate's performance is measured by the number of settings he completes in a specified period of time. The major modifications of this apparatus introduced by the Department of Psychology have been (a) redesign of the mechanical linkage of the stick and rudder to increase its durability and smoothness of operation, (b) substitution of an electronic time-delay relay for a more complicated and less reliable electromagnetic system, and (c) the use of standard replaceable selector switch units.

2. *The S.A.M. Two-Hand Coordination Test.* The S.A.M. Two-Hand Coordination Test is an adaptation of a test purchased from Carl Heinrich who produced it under the supervision of Dr. R. A. McFarland from an earlier model devised by Farmer and Chambers. Studies made under the auspices of the NRC Committee on the Selection of Aircraft Pilots had shown that this test possessed some validity for the selection of successful pilots. Data obtained from the S.A.M. Two-Hand Coordination Test in the Aviation Psychology Program have shown the test to have high validity in the selection of pilots, although not as high as the S.A.M. Complex Coordination Test. It also has some validity for the selection of bombardiers. The test requires the use of both hands simultaneously in manipulating two lathe-type handles to follow a target which moves in an irregular path. The major modifications of the model obtained from Heinrich are (a) the use of a rotating cam to produce a different pattern of movement in the target for each of four one-minute

FIGURE 1—S.A.M. COMPLEX COORDINATION TEST

trials, (b) the redesign of mechanical features of the apparatus to permit mass production and increase dependability and reliability, and (c) the introduction of remote controls for the cycling of test trials and remote recording of the candidate's scores on electric stop-clocks.

3. *S.A.M. Rotary Pursuit Test.* This test is an adaptation of the Koerth Pursuit Rotor which has been used for some years in psychological laboratories.

Some evidence of its validity for the selection of civilian pilots was obtained under the auspices of the National Research Council Committee on the Selection and Training of Aircraft Pilots, and since its use in the Aviation Psychology Program its moderate validity for the selection of pilots and bombardiers has been repeatedly demonstrated. The candidate follows with a hinged stylus a target imbedded in a bakelite disc that revolves at 60 rpm. The candidate's score is the total time the stylus is in contact with the target during fifteen 20-second trials which are separated by 10-second rest periods. The modifications of the Rotary Pursuit Test for mass testing purposes are (a) remote control of the cycling of trials and rest periods and remote scoring of performance on electric stop-clocks, and (b) modification of the hinged stylus to enable standardization of pressure of the stylus on the disc.

At the close of the period of this report a discrimination reaction task was being added to the Rotary Pursuit task in order to introduce the factor of division of attention, on the basis of evidence from PRU No. 1 that this addition increased the validity of the test for the prediction of success in pilot training.

4. *S.A.M. Discrimination Reaction Time Test.* This test was produced in quantity by the Department of Psychology and placed in the classification battery before validation data were available. The candidate is required to react to four different relative positions of a red light and a green light by pushing one of four toggle switches arranged in a square pattern before him. The score is the cumulative total reaction time for eighty reactions. The test has shown moderate but consistent validity for the prediction of success in pilot, bombardier, and navigator training. An improved model of the test, in which all presentations of stimuli during demonstration, practice, and test trials are automatic has been produced recently and supplied to all classification units.

5. *S.A.M. Steadiness-Aiming Test.* The Department of Psychology designed and produced in quantity a Steadiness Aiming Test which attempted to measure the same function measured by the Seashore Photoelectric Aiming Test. The candidate attempts to hold the end of a pivoted and hinged stylus in the center of a small hole to avoid contact with the side of the hole. The test was introduced primarily as a medium for the administration of verbal "stress" stimulation devised by PRU No. 1. The test is now known to have no significant validity for the prediction of training success in pilot, bombardier, or navigator schools either with or without the "stress" stimulation and has been replaced in the classification battery.

6. *Santa Ana Finger Dexterity Test.* This test was designed and produced by PRU No. 3, and introduced in the classification battery as a measure of skill in using the fingers. It is now known to have some validity for the selection of bombardiers. The candidate lifts a small peg, rotates it through 180 degrees and returns the peg to its hole. The score is the number of pegs turned in a fixed period of time.

B. *Psychomotor Research Tests* (*February through July 1942*)

During this period the research tests were used as much to study matters of procedure and details of test administration as to determine their validity for predicting success in training. The tests were given to any available aviation cadets, since the cadets were not at that time receiving regular classification psychomotor tests, and each cadet usu-

ally took all research tests in use at the time. Pilot models of several of the classification tests mentioned above were administered during this period. Tests listed below were administered for validation purposes largely because of their immediate availability.

1. *Lorge Multiple Discriminator.* This test was borrowed from Dr. Irving Lorge, Teachers College, Columbia University, as a representative of a spatial discrimination reaction time test. Five stimulus lights are above five reaction keys and the score is the cumulative reaction time for response to the key below the lamp which is lighted. The test was not suited for mass testing and it showed relatively low validity for the selection of pilots.

2. *Miles Pursuitmeter.* This test has been available since 1922 (3) and was borrowed from Dr. Walter R. Miles, Yale University. The candidate's task is to keep the hand of a meter centered by use of a slide resistor to compensate for erratic movements of the hand induced by an elaborate current variator. Validity coefficients for this test, based on success and failure in primary pilot training were uniformly low, perhaps as a consequence of the difficulty experienced in maintaining the calibration of the test.

3. *McDougall, Rivers, Farmer Dotting Test.* This test was borrowed from Yale University. The candidate attempts to place a stylus in a series of small holes which pass beneath an aperture at an increasing rate of speed. The form of the test used for this research was not suited for mass testing without extensive redesign, and the low validity of the test for pilot selection dictated against such efforts.

4. *Two-Plate Tapping Test.* This test was borrowed from Dr. Irving Lorge, Teachers College, Columbia University, and it was used primarily to obtain a measure of the work decrement. Neither the total tapping score, nor the work decrement score, had validity for pilot selection.

5. *Judgment-Motility Test.* This test was designed by the Department of Psychology and PRU No. 2 at a time when finger dexterity tests were being sought as potential selectors of bombardiers. In the first part of the test the candidate lifts pegs from holes, judges their length, and places them in certain other holes according to length. In the second part of the test the candidate lifts pegs from the holes and returns them to the first holes according to number. In this way each candidate leaves the test ready for the next candidate. The test was administered as a speed test. This test seemed at first to have substantial validity for pilot selection, but further experimentation has shown the validity coefficient to be insignificant.

6. *Bimanual Coordination Test.* This test was devised by PRU No. 1, and used at PRU No. 2 in order to obtain validation data and intercorrelations with other research psychomotor tests at PRU No. 2. The candidate moves a traveler through a triangular pathway by simultaneous manipulation of two handles. The pathway is notched so that the traveler is obstructed when it moves out of the midline pathway. This test has relatively high validity for the selection of pilots, but its relatively high correlations with the S.A.M. Two-Hand Coordination Test and other psychomotor tests mitigated against its use for classification testing.

C. *Psychomotor Research Tests (July 1942 to July 1943)*

During this second period of research testing all tests were administered primarily for validation purposes in a search for new and different

psychomotor tests to replace less valid tests in the classification battery. Subjects were tested immediately after they had finished the psychomotor classification battery, as described in Section A above. Each cadet was given only one experimental test so that the contribution of each test to the psychomotor classification battery could be independently evaluated by use of the multiple regression technique.

1. *S.A.M. Steadiness Test.* This test was supplied to PRU No. 2 by the Department of Psychology primarily for use in the classification battery prior to the completion of the S.A.M. Steadiness-Aiming Test described in Section A above. The candidate attempts to hold a small stylus in an aperture without making contact with the side of the aperture and the score is the total time during which contact with the side of the aperture occurs. The test is now known to have no validity for the selection of pilots or bombardiers.

2. *Two-Hand Steadiness Test.* This test was validated on the hypothesis that it required division of attention. Two copies of the S.A.M. Steadiness Test were used for each candidate and the task was to hold a stylus in the apertures with each hand without touching the sides of the apertures. The test was found to have no validity for the selection of pilots.

3. *Smoothness of Coordination Test.* This test was developed by the Department of Psychology in an attempt to measure the smoothness of the control movements of the candidate in the S.A.M. Two-Hand Coordination Test. The essential mechanism is a pendulum impact recorder which is mounted on the target follower of the test. The total number of impact contacts, which is a measure of the jerkiness of movement, is recorded by an electronic counter system. The scores on this test have very low correlation with the scores on the S.A.M. Two-Hand Coordination Test and with other tests in the classification battery. It was found that the test had no validity for the selection of pilots or bombardiers. It is believed that the principle of the test may be applied to test situations involving other types of coordinative movement with more appropriateness and with perhaps some validity for the selection of aircrew personnel.

4. *Two-Hand Pursuit Test (Thurstone).* This test was designed by Dr. L. L. Thurstone, University of Chicago, for the selection of bombardiers. Copies of the test were produced by the Department of Psychology and validation studies were made at PRU No. 2 and at Ellington Army Air Field. Using the mechanism of the S.A.M. Two-Hand Coordination Test, the candidate attempts to keep a rotating target in the center of a viewing piece which contains a reticle. The candidate is not aware of the fact that the target is moving in a circle since his own movements interfere with the regularity of the apparent motion of the target. This test has been shown to have moderate validity for the selection of pilots and for the selection of bombardiers, but its high correlation with S.A.M. Two-Hand Coordination Test mitigated against its adoption for classification purposes.

5. *S.A.M. Visual Coincidence Reaction Time Test.* The Department of Psychology designed and produced a number of copies of this test in an effort to produce a selection test for low altitude bombardiers. The test was never validated for low altitude bombardiers, but has shown moderate validity for the selection of precision bombardiers. It has no validity for the selection of pilots. The candidate's task is to operate a toggle switch at the instant when a moving bar of light passes between a pair of fixed bars of light, all light being

transmitted through a convex translucent screen. The speed of the moving light and the position of the fixed lights were automatically varied from trial to trial. The "all or none" scoring method on this test resulted in relatively low reliability of scores and further modifications are being made in the test.

6. *Action Potential Index.* Tenseness during flight has been indicated in field studies as a reason for elimination from pilot training. As a preliminary to an attempt to measure muscular tension by a simple method suitable for mass testing, a technique for measuring action potentials in the arm of a candidate who is performing on the S.A.M. Multidimensional Pursuit Test (see

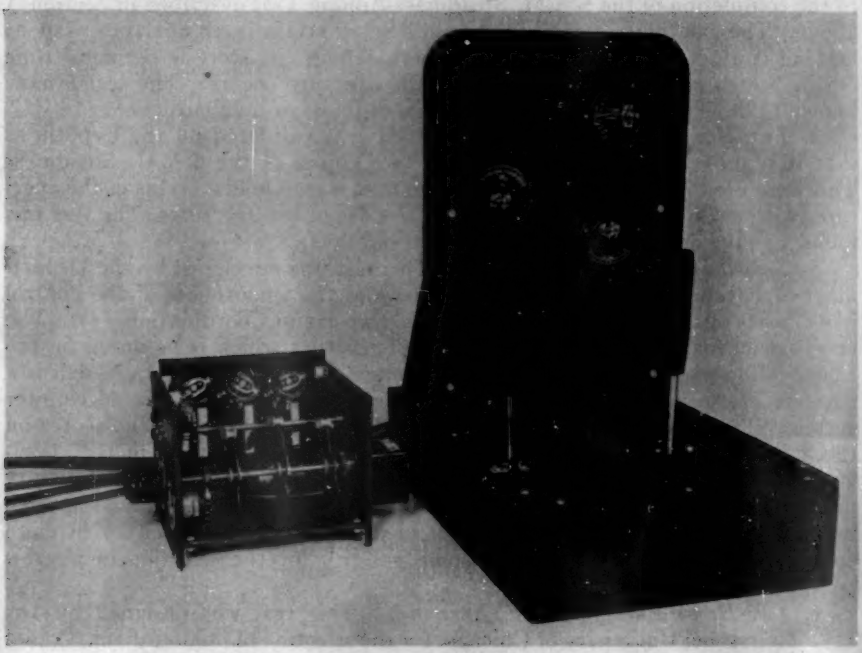

FIGURE 2—S.A.M. MULTIDIMENSIONAL PURSUIT TEST, TABLE MODEL

Figure 2) was developed and administered to a large group of candidates. Validation data are being awaited.

7. *Single Dimension Pursuitmeter.* This test was developed by the Department of Psychology as representative of the simplest form of compensatory pursuitmeter, the task being similar to that of the Miles Pursuitmeter but the mechanism being suitable for mass testing. The candidate moves a wheel, resembling an airplane control, in and out of its shaft in order to hold a horizontal bar in the center of a circular aperture. The position of the bar in the vertical plane is irregularly disturbed by a cam. Validation data are awaited.

8. *Artificial Horizon Pursuitmeter.* This test was developed by the Mechanical Research Laboratory, School of Aviation Medicine, and was validated at PRU No. 2 as representative of two-dimensional compensatory pursuitmeter. The candidate moves a wheel resembling an airplane control in and out on its shaft and rotates it in order to hold a horizontal bar in the center of a circular

aperture and in a horizontal position. The position of the bar in the vertical plane and its deviation from the horizontal plane is irregularly disturbed by cams. Validation data are awaited.

9. *S.A.M. Multidimensional Pursuit Test, Table Model.* This test (Figure 2) was developed as representative of a complex compensatory pursuit test in which movement in three visual objects must be compensated for by movements of a "stick" and "throttle." The candidate attempts to keep the needles of three meters at a zero position by manipulation of the "stick" (in two dimensions) and "throttle" controls while voltage variators disturb the position of the needles in an irregularly manner. This test samples not only eye-hand coordination but also the capacity to attend to a variety of tasks at the same time. Validation data are awaited.

10. *S.A.M. Multidimensional Pursuit Test.* This test is similar to the Multidimensional Pursuit Test, Table Model, described above, except that the candidate must compensate for the movement of the needles in four meters and employs a "rudder" in addition to the "stick" and "throttle." Validation data are awaited.

11. *Stevens Pursuitmeter.* This test was developed by Dr. S. S. Stevens, Harvard University and purchased by the Department of Psychology for modification and validation testing at PRU No. 2. The candidate manipulates a "stick" and "rudder" in order to keep a spot of light in the center of an oscilloscope screen and a needle at a central position on a meter. The task involves, therefore, some division of attention between visual objects, but, it is believed, not as much as is involved in the S.A.M. Multidimensional Pursuit Test. Validation data are awaited.

12. *Electronic Pilot Pursuit Test.* This test was developed by the Electronic Supply Company, Worcester, Massachusetts, and purchased after extensive modifications under the direction of the Department of Psychology. A moving boom carries two photo-cells which must be spanned simultaneously by a bar of light which is controlled by a "stick" and "rudder." The photo-cells move in a vertical direction, a horizontal direction, and also around the axis of the boom. This test differs from all the pursuit tests previously described in that it involves a "following" pursuit function rather than a "compensatory" pursuit function. Validation data are awaited.

13. *Tridimensional Airplane Pursuit Test.* This test was developed by Dr. R. H. Seashore and loaned to the Department of Psychology for use in producing a pilot model of a form of the test adapted to mass production and mass testing. A model airplane rotates on its vertical, horizontal and longitudinal axes in an irregular manner as determined by a set of cams. The task of the candidate is to keep the plane in a straight and level position by manipulating a "stick" and "rudder" control. The School of Aviation Medicine's model of this test has not yet been administered for validation purposes.

IV. Development of Special Techniques for Mass Psychomotor Testing

Starting with relatively unstandardized pieces of laboratory equipment as original test models, a variety of problems in apparatus design and construction and in testing conditions is being met in developing satisfactory procedures for the daily administration of standard tests to thousands of aircrew candidates at different stations throughout the

UNIVERSITY OF MICHIGAN LIBRARIES

country. Those concerning apparatus have been almost entirely handled by the Department of Psychology, while PRU No. 2 has been concerned with conditions of testing. However, all Units and many individuals in the Aviation Psychology Program have contributed to the solution of these problems.

A. It was first determined to construct all classification tests in units of four pieces of test apparatus and one central control unit. It was found from early investigation that one examiner could give instructions, record scores, and proctor the testing of four subjects at one time. Fewer or larger numbers of subjects in one test room resulted in inefficient use of personnel.

B. Preliminary testing indicated that an adequate behavior sample, yielding satisfactory reliability of measurements, could be obtained for each of the classification tests in thirteen minutes of over-all testing time, of which about one-third is spent in instructions, demonstrations and practice, and the remaining time in testing for the record. Two minutes are allowed for movement between test rooms. Later studies have shown that the validity of the present classification tests is not increased when longer tests are given. The control mechanisms developed by the Department of Psychology administer the tests so that they automatically keep up to this schedule.

C. All control equipment has been designed to yield maximum automaticity of operation, so that the test examiner is left free to proctor and to record scores accurately. An essential principle has been the consolidation of all control switches and scoring devices on a control desk.

D. In the choice of apparatus designs and parts at least two important considerations have governed. (1) Ruggedness of construction has been essential in order for equipment to withstand continuous daily use. The substitution of castings for pieced framework, and of gears and sprockets for cables and friction discs are typical examples. (2) Procurement difficulties have often made it necessary to choose standard mechanical and electrical parts in terms of their availability with the priorities obtainable rather than in terms of ideal characteristics.

E. Problems of maintenance in the field have been met in two ways: (1) the use, wherever possible, of instantaneously replaceable and interchangeable components of apparatus; and (2) the preparation of schedules for regular preventive maintenance check-ups and standardized instructions for necessary repairs in the field. These have been prepared by the Department of Psychology on the basis of maintenance records from PRU No. 2 and other units.

F. The development of adequate standard instructions for each test has been a major concern of PRU No. 2. Studies were made of the fre-

quency of misunderstandings of instructions by candidates. Instructions have been successively revised and tabular studies again made until maladaptations as a function of inadequate instructions or demonstrations have been reduced to a minimum. Examiners have consistently been required to repeat all instructions verbatim from memory, to give demonstrations only in a prescribed manner, and to answer candidates' questions in standard fashion. A comprehensive system of psychomotor examiner training and systematic checking has been developed. Analysis of variance studies have shown that variance attributable to the test examiner has become insignificant.

G. Extensive analysis of variance studies have shown that variance related to the apparatus on which the subject is tested is highly significant, even though pieces of apparatus have been produced by commercial firms with precision mass production methods from models and complete drawings. Until more accurate machining and more intensive calibration procedures are accomplished, it is necessary to control these differences statistically. As routine practice all scores obtained on each piece of apparatus are tabulated, means and standard deviations for successive groups are computed, and conversion tables are revised whenever means or variabilities of scores consistently show significant deviations.

BIBLIOGRAPHY

1. GLENN, C. R. A preliminary report on a performance test for flying. *J. Aviat. Med.*, 1935, **6**, 14–19.
2. MASHBURN, N. C. Mashburn automatic serial action apparatus for detecting flying aptitude. *J. Aviat. Med.*, 1934, **5**, 155–160.
3. MILES, W. R. The Pursuitmeter. *J. exp. Psychol.*, 1921, **4**, 77–105.
4. Staff of the Psychological Branch, Office of the Air Surgeon, Headquarters Army Air Forces. The Aviation Psychology Program of the Army Air Forces. *Psychol. Bull.*, 1943, **40**, 759–769.
5. Staff, Psychological Research Unit No. 1. History, organization and procedures Psychological Research Unit No. 1, Army Air Forces. *Psychol. Bull.*, 1944, **41**, 103–114.
6. Staff, Psychological Research Unit No. 3. Organization and research activities Psychological Research Unit No. 3, Army Air Forces. *Psychol. Bull.*, 1944, **41**, 237–245.
7. Technical Manual, TM 8-320. Notes on psychology and personality studies in aviation medicine. Sup't of Documents, Washington, D. C., Jan. 27, 1941.

UNIVERSITY OF MICHIGAN LIBRARIES

A BLINKER ADAPTATION OF THURSTONE'S 1943 CODE APTITUDE TEST

HILDING B. CARLSON AND JOSEPH RICH
University of Illinois

This report* is the outcome of a request by Lt. Commander Charles H. Bowman, commanding officer of the Naval Training Schools at the University of Illinois, for assistance in selecting or devising a test that would enable the Signal School to predict the probable success of its trainees. Approximately 25 per cent of an average class fail to pass the 16 weeks' training period in the Signal School due to inaptness. This figure is somewhat below that reported by Taylor (3), who states that from 30 to 60 per cent of unselected men who begin code study in commercial and military schools fail to become proficient operators. Without adequate selection, therefore, considerable instructor and student time are lost in the attempt to train the potential failures. Our problem was to select or devise a test that would differentiate those who are more likely from those who are less likely to be successful in the Signal School.

Trainees at the Navy Signal School are required to learn semaphore and flag hoist signaling as well as Morse code using blinker lights. Preliminary investigation of the conditions under which the men are trained and the difficulties they encounter, in so far as they were known to the instructors, revealed that the most difficult task is to learn to receive the blinker code signals. It was therefore decided to investigate this aspect of the problem first.

The speed at which blinker code signals must be received for successful completion of the 16 weeks' course is slower than that for auditory code signals. Blinker trainees who are unable to receive at a rate of at least seven words of five characters each per minute are bilged (failed). Trainees who are able to receive at a rate faster than this but less than

* Numerous people have aided us in this project. Lt. Russell I. Damon, Executive Officer of the school, has given extensively of his time and assistance. Without Lt. Damon's aid, the work could not have been carried on successfully. Lt. C. G. Browne, former Personnel Officer of the Navy Signal School offered frequent advice and considerable assistance throughout the early stages of the work. Mr. Albert James Ebel, Chief Engineer of Radio Station WILL, designed the dual selective amplifier described in a succeeding article. He also gave considerable time to the cutting of the test records. Chief Petty Officer Frank A. Simmons of the Navy Signal School worked with us for several months, both in the manual administrations of the test and in cutting of the records. Earlier manual presentations were given with the help of Signalman Second Class Morris Vorenberg. Miss Alberta Menzel of the staff of Radio Station WILL read the instructions for the recordings. Professor Herbert Woodrow has given us encouragement and financial support throughout the project. Since the test reported here is but a blinker modification of Professor L. L. Thurstone's Code Aptitude Test, our indebtedness to him is apparent.

ten words per minute are graduated as Seaman First or Second Class, while those who are able to receive at a rate faster than ten words per minute are given the higher rating of Signalman Third Class. The problem, therefore, can be more precisely stated as being the selecting or devising of a test that would differentiate those trainees who have the possibility of learning to receive the blinker code signals at a rate of at least 35 characters per minute from those who do not.

We have been unable to locate any published report of a visually presented code aptitude test. Both Army and Navy have been using Thurstone's 1919 autitory rhythm test (4) which consists of a series of code letters and code letter groups presented auditorially in pairs which the subjects are to identify as the "same" or "different." Thurstone (5) reported recently on a work sample test having considerably higher reliability and validity than the rhythm test alone.* Since there probably is a fairly high correlation between an auditorially and a visually presented test, one possibility would be to make direct use of Thurstone's new test. However, by presenting a work sample test using the regular blinker signal lights it is possible to approximate more closely the conditions the trainees face during their training period as well as when on board ship. Further, some instructors report that they have had more difficulty training men to receive blinker signals when they have had previous auditory code experience than in training those with no previous code experience at all.

For these reasons, our first attempt consisted in modifying Thurstone's work sample test by presenting the code signals visually. The second edition of Thurstone's work sample test (6) consisted of 14 subtests, namely: 1) an auditory rhythm test; 2) letter D; 3) letter R; 4) letters D, R; 5) letter U; 6) letters DU and RU; 7) letters D, R, and U; 8) letter S; 9) letters D, R, U, and S; 10) auditory distractors with letters D, R, U, and S; 11) recognition of letter D, R, U, or S interspersed among other signals; 12) auditory number span; 13) auditory rapid spelling; and 14) dot discrimination similar to the rhythm test. Discussions with code instructors indicated that the last fifteen items of the auditory rhythm test probably would be too difficult when presented visually, so these items were omitted. Since the maximum speed for blinker work is only about 17 words per minute with the receiver working with a partner, there seemed to be little need for the rapid spelling test. Since the modified test was to be presented visually, the auditory memory span subtest was likewise dropped.

Inasmuch as the trainees in the testing situation would have to write their own responses, and since they could not write effectively while looking at the blinker lights, it was necessary to allow them more time between test items than in the auditory situation. After some prelim-

* Since this project was started, D. W. Taylor reports a similar test and concludes that "an Initial Learning Test seems to be the most promising test yet developed for use in the selection of men for training in code." *Learning Telegraphic Code. Psychol. Bull.*, 1943, 40 (July).

UNIVERSITY OF MICHIGAN LIBRARIES

inary study, the intervals between test items were set at from three to five seconds, depending on the number of letters in each item that the subject had to record. This interval was found to be long enough to enable them to record their responses, but not long enough for them to become bored.

In the interference subtest both visual and auditory distractors were used in the first administration. The interferences consisted of 1) a man walking across the stage platform, 2) a ruler clapping on the floor, 3) two men conversing, 4) the overhead stage lights turning on and off at an irregular rate, and 5) the footlights flashing on and off at an irregular rate.

A few other minor changes had to be made. All instructions referring to auditory functions, such as "listen," were changed to visual terminology, such as "watch" and "look."

Test Situation. In the test situation, two blinker signal lights were located about four feet apart on a table in the center of the stage in the testing room. These blinker lights are ship's lanterns which hang four and one-half feet apart on a yardarm during training. Each light contains three or four type T7 Hygrade, 120 volts, 15 watt bulbs. Surrounding the bulbs is a magnifying lens. The code operator was stationed in a wing, as far from the trainees as possible, in order to minimize the sound of the telegraph key. The instructions were read by the junior author from a position on the stage intermediate to the blinker lights and the code operator.

Subjects. Of the men in the first class studied at the Naval Training School (Signal) at the University of Illinois, only 92.0%* were available on March 29 to take the first manually presented modified form of Thurstone's second experimental test before their training period began. This 92% group will be considered to be the entire experimental group.

Since the purpose of this experimental administration was to determine the differentiating power of each subtest as well as of the test as a whole for future groups of trainees, the representativeness of the class studied is an important consideration. The relevant data that were available for the members of this class and for those in the earlier classes with which this class was compared, include average age, average schooling, average scores on the general classification test, average scores on Thurstone's 1919 auditory rhythm test, the first, second, and third choices of specialization within the Navy, and the percentage of those who failed from the Navy Signal School. However, since no measures of scatter were kept for the first classes, only crude comparisons could be made between the first classes and the class studied. For the most part, the mean scores for the two groups are not markedly dissimilar. Insofar as the mean scores are concerned, therefore, the members of this class were considered as typical of the trainees at the Navy Signal School.

Scoring. Each subtest was separately scored for errors, the scoring facilitated by the use of stencils. A unit error was defined as a failure to respond or an incorrect letter whether that letter appeared alone or as part of a group of letters forming the test item.

* At the request of the Navy, all figures are given in terms of percentage of the total group studied.

Distribution of Errors. Error distribution curves of each of the subtests were drawn. Inspection of these graphs revealed that on three of the subtests there were many perfect scores and few errors by any trainee. These subtests—the second, third, and fifth—are primarily training exercises involving recognition of the position of the signal for letter D, R, or U when it was interspersed with two other signals. The distributions indicate that these three subtests could have very little discrimination value and the scores on them were therefore disregarded. This reduced the number of subtests included in the total score to nine. Table 1 shows the distribution of the error scores on the test as a whole, each score being the sum of the errors on the nine subtests. The least number of errors is four, the highest number 536, and the median 74.5.

TABLE 1

DISTRIBUTION OF ERRORS IN TERMS OF PERCENTAGE OF TRAINEES IN EACH CLASS INTERVAL, BASED ON SUM OF NINE SUBTESTS

Class Interval	Per cent
0– 30	13.0
31– 60	26.8
61– 90	21.6
91–120	10.2
121–150	11.5
151–180	7.7
181–210	3.7
211–240	2.0
241–270	1.5
271–300	0.0
301–330	1.0
331–360	0.5
....	
511–540	0.5
Total	100.0

Validity. The administration and the manual scoring time of this test were considerable. It was desirable to reduce either or both without lowering the validity or reliability. With a view toward eliminating those subtests with low validity, validity coefficients were determined for each subtest. The criterion first used was continuance in or bilging from the training school after the first six weeks period. This particular period was chosen because it was the time at which the first group of bilgees was dropped. This criterion necessitated the use of bi-serial or tetrachoric correlational procedures. Inasmuch as both time and clerical help were limited, tetrachoric correlations only were determined (2). The resulting correlations are included in Table 2.

UNIVERSITY OF MICHIGAN LIBRARIES

TABLE 2

VALIDITY COEFFICIENTS (TETRACHORIC) OF EACH SUBTEST AND OF TOTAL TEST WITH CONTINUANCE IN, OR BILGING FROM CLASS AS CRITERION

Subtest	Validity Coefficient after 6 weeks	after 12 weeks	after 16 weeks
Rhythm	.48		
Letters D & R	.53		
Letters D, R, U	.48		
Letters D, R, U, S	.76		
Interference	.68		
Letter Recognition	.23		
Letters DU & RU	.35		
Letter S	.33		
Dot Signals	.40		
Total, Nine Tests	.73	.54	.46
Total, First Six Tests	.68	.59	.51
Per cent of trainees in each group	100	93.3	93.3

The subtests with lowest validity are letter recognition, letter S, letters DU–RU, and dot discrimination, in the order named. It was decided to drop the scoring for letter S and letter DU–RU but to retain them in the administration of the test since both give training for later, more differentiating subtests. Both the dot discrimination and the letter recognition tests might be eliminated entirely. Since the dot discrimination test is somewhat similar to the rhythm subtest, the decision was made, more or less arbitrarily, to disregard the scores on this subtest and to drop it from future administrations. The decision to retain the letter recognition test was also more or less arbitrary, and cannot be completely justified at present.

Omitting scoring of the subtests for letter S, letters DU–RU, and dropping the dot discrimination subtest, reduces the number of subtests to be scored to six. Error scores for the total of these six subtests were then calculated. The median error score is then 56.8. The least number of errors is three and the largest number is 293.

The validity of the blinker code aptitude test, based on six subtests, and calculated in the same manner as for the several subtests, gives coefficients of .68 at the end of the six weeks period, .59 at the end of the 12 weeks period, and .51 at the close of the training period. Comparable figures for the nine subtests gives .73, .54, and .46 respectively. Comparing these values with those for six subtest, indicates that little or nothing has been lost by the exclusion of three subtests.

As noted in Table 2, N varies from 100% of trainees tested for the first six weeks to 93.3% for the twelfth week. After the six weeks period, 6.7% of the men were discharged from the class studied for various

reasons: to return to Coast Guard units, to enter V-12 training program, for sickness, and for discipline. Since these trainees left for reasons other than inaptness in code, they cannot be considered as being successful or as having bilged.

Of the trainees who were not transferred out of the school, 20.1% reported that they had had some code training of one kind or another prior to the time of the test. Since this previous training may influence the validity coefficients unduly, the records of these men were withdrawn and tetrachoric coefficients calculated for those who did not report having had code training. The validity coefficients, based on the sum of the scores for the six subtests for the trainees who did not report having previous code training is .51, while on the basis of nine subtests it is .52.

As stated earlier, the precise purpose of this experiment was to select or devise a code aptitude test that would differentiate those who could learn to receive blinker code signals from those who could not do so, under the training procedures at the Navy Signal School. The criterion of continuing in and graduating from versus bilging out of the Navy Signal School involves other factors than learning to receive blinker signals. A more appropriate criterion would describe more accurately the effectiveness of the test. One such criterion is the class performance while learning to receive blinker signals as graded by the instructors. Accordingly, the final blinker grade in the course was used as a validating measure. This final blinker grade is equally dependent upon the weekly averages and upon a final qualifying examination, both weekly grades and final examination grade being determined by instructors. For men who bilged out, the average weekly grade up to the time of bilging was used, since this included all the data that was available. Tetrachoric correlations between the average class performance grade thus determined, and the code aptitude test, for both six and nine subtests, using all available trainees, are .64. When only those trainees who did not report prior code training are included, the validity coefficient (tetrachoric), with average class performance as the criterion, for the six subtests is .66.

A validity coefficient that high clearly indicates that the test is a useful one for the purpose of selecting blinker signalmen.

The validity coefficients for each subtest, with average blinker grade as the criterion are as follows: rhythm .43; letters D & R, .64; letters D, R, U, .61; letters D, R, U, S, .63; interference, .54; letter recognition, .42; letters DU & RU, .48; letter S, .27; dot signals, .52. These coefficients give further evidence for the justification for including in the final scoring subtests D & R, DRU, DRUS, and the interference subtest. They also give additional evidence for not scoring subtests S and DU & RU. There is some question as to whether the rhythm or the dot

UNIVERSITY OF MICHIGAN LIBRARIES

discrimination subtests ought to have been retained, and the question can even be raised as to whether both might not be dropped in the effort to cut down the administration time. Unless further evidence to the contrary is obtained, the letter recognition subtest might also be dropped entirely. These possibilities will be considered when data on the recorded edition of the test is available.

A method other than the correlational procedure to show the validity of the test is to indicate, in a distribution curve, the scores of the bilgees. This has been done in Figure 1. Error scores in class intervals of

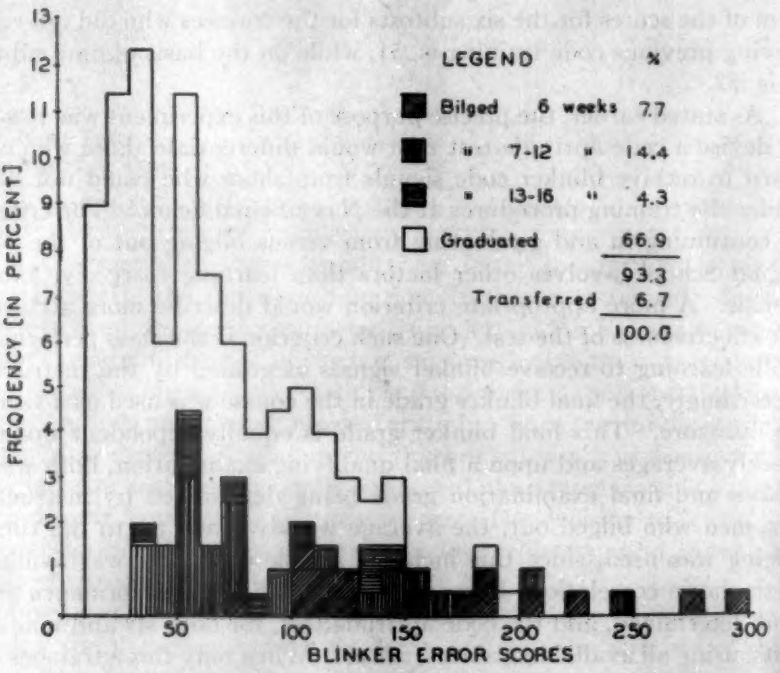

FIGURE 1—DISTRIBUTION OF ERROR SCORES ON BLINKER CODE APTITUDE TEST (SIX SUBTESTS). FIRST CLASS STUDIED

ten for the sum of the six subtests selected are indicated on the abscissa and the frequency of the scores on the ordinate. The total graph area includes 93.3% of the men tested. The areas with diagonal lines represent those who bilged at the end of the six weeks period, the areas with vertical lines those who bilged between the seventh and twelfth weeks, and the horizontal lines those who failed between the thirteenth and sixteenth weeks.

There is a third method by which the validity of the test might be indicated. This method, for the present purposes, may be of greater practical value than either the correlational procedure or the graphic

technique. This method would be to show the percentage of failures among those who were not transferred for other reasons who had test scores in the lowest five per cent, using the sum of the six subtests selected, how many in the lowest ten per cent, etc. The results in terms of the per cent of trainees who bilge below various centile levels is as follows:

Below centile 5,	80%
Below centile 10,	75%
Below centile 15,	67%
Below centile 20,	60%
Below centile 30,	50%
Below centile 35,	47%
Below centile 45,	43%
Below centile 65,	40%

It is not the function of this report to determine a critical test score below which men might well be excluded from the Navy Signal School. If we can assume that this class is representative of future classes, then those men who score in the lowest 5% have only 1 chance in 5 of graduating; those in the lowest 10%, have 1 chance in 4 of graduating; those in the lowest 30% have 1 chance in 2. On the other hand, those who have scores in the highest 25% are almost certain to be successful. If some critical score were to be established, and the test administered at induction centers, considerable time, money, and effort could be saved.

Reliability. The reliability of each subtest and of the test as a whole was determined by means of the split-half method, making use of the Spearman-Brown estimation formula. The estimated reliabilities thus obtained for each subtest and for the test as a whole are:

Rhythm	.79
Letters D, R	.69
Letters D, R, U	.90
Letters D, R, U, S	.92
Interference	.94
Letter recognition	.73
Total of Six Subtests	.97

While test-retest reliabilities probably would be more accurate, and give somewhat lower values, those obtained are so high that there can be little question that the consistency of the test as a whole is very high.

Additional data from second class studied. Since the preparation of this report, additional data on the validity of the blinker code aptitude test was obtained from a second class. Ninety-seven and one-half per cent of the men in this class took the test in April. At the end of the training period only seven per cent had failed, indicating that this class was considerably superior to most classes.

The distribution of the summed errors on the six subtests selected for scoring for the second class studied is given in Figure 2. These subtests are the first six listed in Table 2. The disposition of the trainees

UNIVERSITY OF MICHIGAN LIBRARIES

after their 16 weeks of training is also indicated therein. The men who bilged are signified by diagonal shading, Seaman Second Class by vertical lines, Seaman First Class by horizontal lines, and Signalman Third Class by blank spaces. Median error scores for the men in each of these

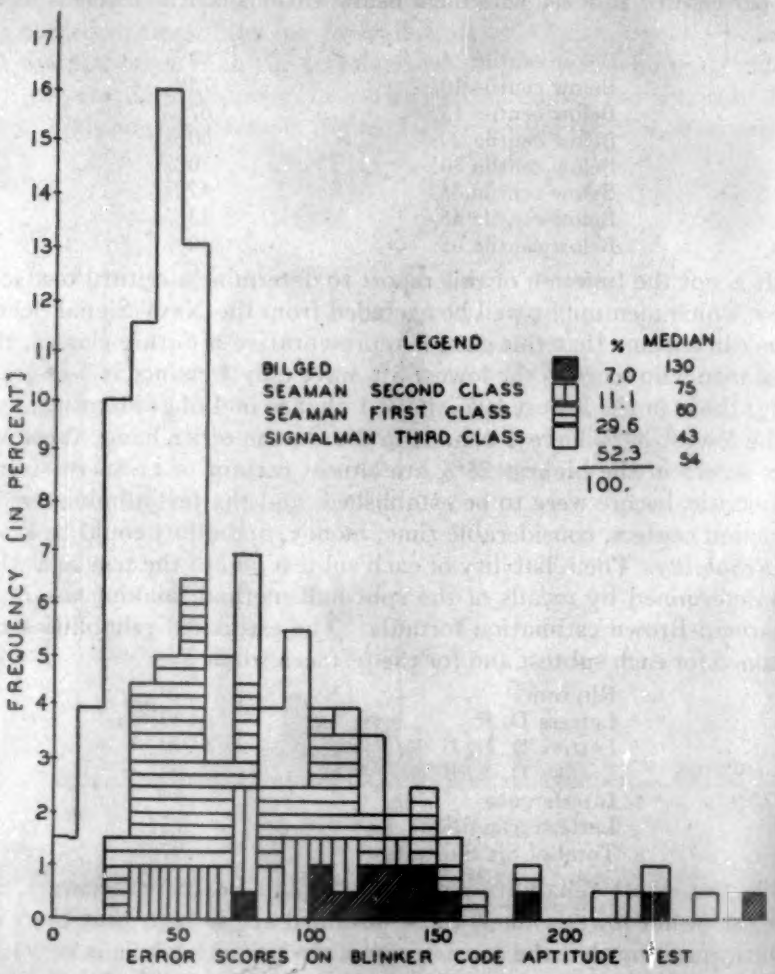

FIGURE 2—DISTRIBUTION OF ERRORS AND DISPOSITION OF TRAINEES AFTER 16 WEEKS OF TRAINING. SECOND CLASS STUDIED.

four classifications are 130, 75, 60, and 43 respectively. While there is considerable overlapping in error scores between the men in these classifications, the median error scores indicate clearly that the test has considerable differentiating power. If a critical score of 100 had been adopted, eliminating from signal training some 22 per cent of those sent

to the Navy Signal School, all but one of the failures would have been excluded immediately.

The validity of the test, using the data for the second class and tetrachoric correlational procedures, with graduating-bilging as the criterion, is approximately .65. This value, however, should not be given much weight, since its determination had to be made with the division of the criterion near the tail of the distribution. With final blinker code grades as the criterion, the validity coefficient is .63. In this case the division of both distributions was made near the median scores and this value can be given considerable weight. Inasmuch as these values are similar to those reported for the first class, they substantiate the high validity already reported.

Recorded edition of the blinker code aptitude test. On the basis of data obtained in the administration of the test to groups of trainees in March, April, May, and June, and with a view toward reducing the time consumed in administration, certain modifications have been made in the test as first administered in March. With these modifications, test records were cut in July to be used in conjunction with a dual selective amplifier (1). The first experimental recorded edition of the test include the following subtest: 1) rhythm test; 2) letter D; 3) letter R; 4) letters D, R; 5) letter U; 6) letters DU and RU; 7) letters D, R, and U; 8) letter S; 9) letters D, R, U, and S; 10) auditory distractors with letters D, R, U, and S; and 11) letter recognition. Of these, numbers 2, 3, 5, 6, and 8 are not scored. By means of the phonograph records, the test was first presented in July to another class of men in a more standardized manner than is possible by manual means. A follow-up of this group will be reported when data are available.

Summary. An experimental blinker modification of Thurstone's 1943 code aptitude test has been administered to two classes at the Naval Training School (Signal) at the University of Illinois. The test has been demonstrated to have both high validity and reliability, thereby showing that it can be a useful device for the prediction of success and failure in the learning of the blinker code.

BIBLIOGRAPHY

1. Carlson, H. B., and Ebel, A. J. A dual selective amplifier. (To appear.)
2. Chesire, Leone, Saffir, M., and Thurstone, L. L. Computing diagrams for the tetrachoric correlation coefficient. Univ. Chicago Bookstore. 1933.
3. Taylor, D. W. Learning Telegraphic Code. *Psychological Bulletin*. 1943, 40, 461–487.
4. Thurstone, L. L. Mental Tests for prospective telegraphers. *J. Appl. Psychol.*, 1919, 3, 110–118.
5. Thurstone, L. L. Preliminary report of a code aptitude test. Privately printed. February, 1943.
6. Thurstone, L. L. Report on a code aptitude test. Privately printed. May, 1943.

UNIVERSITY OF MICHIGAN LIBRARIES

MASSERMAN, JULES H. Neurosis and behavior. Chicago: University of Chicago Press, 1943. Pp. xv + 269.

The author states in his introduction that it is his purpose "in the following pages to show that the dynamic concepts of inner needs, holistic adaptation, equated or symbolic expression, and motivational conflicts—concepts derived from psychoanalytic theory—can be broadened and profitably applied to the experimental investigation of animal as well as human behavior."

A first series of experiments on cats studied the role of the hypothalamus in emotion. By direct stimulation through a bipolar needle electrode the symptoms of sham rage were produced. This was distinguished from "true affective states" by being irrelevant to the external situation, failing to persist when the stimulus was discontinued, and several other criteria. This sham rage could not be conditioned to light, sound or air-blasts in many trials. Though the author points out that the behavior is initiated by the direct stimulation of efferent systems, his own interpretation is not that association must involve afferent or connector systems, but rather that sham-rage lacks motivational connotation or "meaning" and therefore can neither be conditioned nor made the basis for an experimental neurosis (58).

A second group of experiments dealt with a conditioned feeding response, using a light flash and bell as the signal of the manual or automatic delivery of food to a food-box in a cage. These experiments should be of particular interest to psychologists not only because they are well observed and well described, but because Masserman has not fallen into the prevalent stultifying practice of observing and recording only totals and averages of many trials. He is more interested in learning than in learning curves, and, when the returns justify it, gives a play by play account which throws light on the nature of learning which is entirely lacking in most reports of maze-running or of long series of conditioning trials.

The use of a switch-lever which operated the flash and bell signals was learned by the cats under various circumstances. Masserman notes the use of different methods for raising the lid of the food-box or for depressing the lever. "Cats which learned to depress the switch with the paw (61) generally used that motion thereafter; those which learned by accidentally rubbing against the switch continued to rub against it, usually with the head, occasionally with other parts of the body . . . ." There were some cats which displayed eagerness to get into the experimental cage and resistance to being removed. The description of results is used to point up Masserman's contention that the Pavlovian conception of the reflex misses much significant behavior which he believes can be understood only in terms of desire, motive, need, sign gestalt, etc.

In these experiments in conditioning barring the passage to the food-box did not bring about neurotic changes. "External frustration" in general produced adaptive extinction of the response. But results were very different when "the goal-seeking behavior . . . was made conflictful by the elicitation of opposing motivations within the animal itself." This was achieved by subjecting the cat either to an air-blast or to a grid shock at the moment of feeding. In response to the air-blast, animals usually rushed to crouch at the far side of the cage. In response to shock, animals gave a startled jump and then with slow dignity stalked away from the food after the shock had ended. In some instances only one such experience was enough to produce in the cat severe

neurotic symptoms. Food signals "which formerly had occasioned an immediate rush to the food-box, now induced crouching, hiding, vigorous attempts to escape from the apparatus, rapid and irregular pulse and respiration, trembling, and other behavior suggestive of marked anxiety or even panic (68)." In most animals these phobic reactions persisted almost unchanged for months without further reinforcement by shock or blast (68).

Certain animals developed "counterphobic" patterns, stereotyped mannerisms which constituted a defensive adaptation by reducing tension. Regressive behavior in the form of "self-preoccupation" and "autistic indulgence" (cleaning and licking) are interpreted by the author as substitute satisfaction for the frustrated drives.

The neurotic symptoms were in general increased when hunger was intensified. They were also increased by pushing the animal toward the food-box, especially immediately after feeding signals. This produced symptoms of extreme claustrophobia and uncontrollable panic.

Rest left the abnormal behavior unchanged in 37 animals; in seven it produced some amelioration. Positive therapeutic measures, however, were found to be distinctly effective. These included feeding the animal manually or even forcibly, just before placing it in the cage. The suggested explanation is that one of the conflicting motives had been reduced.

In other animals, handling, stroking and petting gradually enabled the animal to resume feeding at the signal, but this was often stopped when the stroking was discontinued.

In still other animals the intensification of hunger reestablished eating in spite of conflict, and this initiated the reduction of the conflict. Hunger could not be intensified through starvation because animals would go for from eight to twenty-two days without eating though food was available in the food-box. The intensification was brought about by the use of the sight or smell of food, or by the use of sensory stimuli associated with satisfactory feeding or by forcing the cat into closer proximity to the food.

Another "therapeutic" measure consisted in placing a normal cat in the cage with the neurotic animal. This had in a few cases a slight effect toward restoring normal behavior in the neurotic.

A most interesting final therapeutic measure was to train the animal in advance to depress a switch which actuated the signals for reward. This switch was for a time severely avoided after the blast or shock; but such animals tended generally to resume the use of the switch and to manipulate it with increasing confidence until they had reestablished their feeding routine. This method was the most effective of all the measures.

Following the accounts of his experiments is an extended survey of the literature on conditioning and neuroses, collected as evidence of the author's contention that the concepts of association and Pavlovian conditioning are inadequate for the description of such behavior. He himself offers four "fundamental principles of behavior."

1. Behavior is motivated by the biologic needs of the organism.

2. Behavior is contingent upon, and adaptive, to the *meanings* of the "objective" and social environment as interpreted by the individual organism.

3. Behavior relieves bodily tensions not only by direct, but also by *substitutive* or *symbolic*, activity.

4. When psychobiologic motivations or environmental meanings become excessively confused or conflictful, behavior correspondingly becomes abnormally indecisive, sub-

UNIVERSITY OF MICHIGAN LIBRARIES

stitutive, diffusely symbolic, and biologically inefficient, i.e., "neurotic" or "psychotic" in expression.

Carmichael, in a delightful introduction to a recent *Bulletin* article on the "Experimental Embryology of Mind" asks at what point the fertilized ovum, until then the creature of physics and chemistry, is supposed to become independent of its origins and acquires a soul and is ruled by its destiny. The answer is, of course, that your real vitalist objects to explanations in terms of proximate cause at any point. Even the sperm and ovum are moving on toward their destinies.

Many psychologists that the author of this book would class as mechanists can sympathize heartily with his objections to the crude notions of Pavlov's reflexology and the uncritical acceptance of those notions by some American experimenters. But he would find if he actually made the attempt, which he has surely never done, that many of his own reported facts take very kindly to description in terms of the circumstances leading up to rather than the result led toward. There is no conflict between final cause and proximate cause. The same event may be given both descriptions. "Mechanists" do not deny meaning and purpose in the world. They are attempting to describe meaningful and purposeful behavior in terms of the circumstances under which it occurs, instead of in terms of its expected outcome.

EDWIN R. GUTHRIE

University of Washington

JONES, HAROLD E. Development in adolescence. New York: D. Appleton-Century Company, 1943. Pp. xii + 166.

This book is a report of observations and measurements made on a boy and his environment from his tenth to his eighteenth year of life. The case, alias John Sanders, is one selected from the 200-odd subjects of the California Adolescent Growth Study as exhibiting a variety of problems of adjustment and is presented "as a teaching vehicle—to illustrate method and to present specific growth phenomena concretely rather than abstractly." The data concerning the case were collected by the Institute Staff at intervals over a seven year period.

The method of presenting the growth phenomena is difficult to describe. It is not biographical in the strict sense of being undisciplined artistry; yet there is presented the social setting and, to some extent, the stream of life in which the case developed. The reader at times becomes engrossed with cultural factors: general rather than specific effects of home conditions, school clubs, unattainable ambitions created by our society, and the like. The chapters are somewhat like a number of short stories appearing in one volume; some sequent, some not. The author states that the book treats "specific modes of expression of a total growth process."

One gets a general picture of John Sanders and his environment from the first two chapters. The boy grew up in a section of a city interstitial to the industrial and the better residential districts. Therein he and his associates were neither very well-to-do nor extremely poor. Most of his time was spent close at home under the continual supervision and dominance of his mother. He became a docile, well trained child who gave little trouble. This docile

attitude was carried with him to elementary school where he made little impression on either teachers or pupils. After graduating from a local elementary school, he transferred to a junior high school to become enmeshed in a crowded institution composed of children from socio-economic groups both above and below his own. John rather retained his stay-at-home role and so made few friends with his classmates and none with his teachers. He did become interested in the social activities provided at the Adolescent Growth Study Clubhouse and spent all available time there. He joined the Boy Scouts wherein his role was one of doing the hard work but not one of gaining positions of honor. Upon graduating from junior high school he entered a senior high which was twice as large and which had a wider variety of social make-up. He did not possess the requirements of the social clubs and so had little to do in running either the school or the social life of the students. As a substitute, he participated in clubhouse activities, joined a current events club, and remained a Boy Scout until in the twelfth grade he dropped it for the Y.M.C.A. He was regarded as a queer fellow by most students and as a poor college prospect by most teachers.

Views of special features of John are given in Chapters III through IX. The different features selected for view with the sources of information are presented in different chapters, usually against a group background. These include: academic and social attributes; health, and physical and physiological changes; strength and motor abilities; intelligence, achievement, and motor, verbal, and rational learning abilities; interests and attitudes; needs and drives; and self-evaluations.

Interrelationships of different features of John are discussed in the final chapter. Special stress is placed on the relation of changing physical characteristics to difficulties encountered in social, economic, and general environmental situations. Many of the social difficulties seemed to have a core in John's physical weakness and retarded maturational status. Improvement in adjustment coincided with his final physical maturation and with changes in the group itself as its aspirations became more sober.

The author achieved his aim of presenting specific aspects of growth phenomena concretely and did so in an interesting fashion. The effect of this achievement on various readers and hence its over-all contribution to psychology should be considered. The reviewer is of the opinion that a layman or a novice in psychology will accept the account as either fact or fancy depending upon whether he is already inclined toward determinism or magic. It is improbable that the book would induce him to ask questions or formulate hypotheses as would the trained person when considering single cases. This brings up the more fundamental question of the treatment in psychology of the unique case. The traditional approach which seeks laws or principles regarding the operation of processes is skeptical of it as the apparent relationship between various features is not subject to empirical verification. The dominant mother, for example, appeared as an adverse influence on John Sanders, but one can wonder how many dominant mothers have brought up dominant sons. So can one question the seeming effects on development of the various cultural factors mentioned. The very fact that such questions do arise, however, gives the book its value.

JOHN B. WOLFE

University of Mississippi

UNIVERSITY OF MICHIGAN LIBRARIES

BILLS, A. G. The Psychology of Efficiency. New York: Harper & Bros., 1943. Pp. xiv + 361.

This book brings together information from various sources, as it applies to adjustment from the point of view of efficiency. It deliberately avoids emphasis upon the emotional aspects of behavior; it goes beyond the management point of view by being concerned with the "mental worker himself who would like to know how to accomplish the most with the least wear and tear and the greatest long-time satisfaction to himself." The chief topics covered, in 18 chapters, include: the thinking process; the nature of, measurement of, and recovery from, fatigue; sleep; attention and distraction; the relation of suggestion, learning, motivation, physical conditions, etc., to work efficiency; self-analysis and improvement of work procedures.

It is difficult to determine whether the book is intended as a textbook, as a reference book for professional psychologists, or as guidance for the layman. The language is sometimes that used in the "bend-down" type of popularizing (e.g., frequent use of a body-machine analogy, emotions discussed as "friction," rules for "lubricating" personal human relationships), yet there is probably too much technical material for easy reading by a layman. On the other hand, much of the writing is not precise enough for professional training or reference use. The latter point may be illustrated in several ways. For example, such ambiguous terms as 'tense mental attitude' are employed. There are rather literal physiological treatments of phenomena which are as yet in the realm of constructs (brain energy levels in Ch. II; mental blocks in Ch. IV). There is some uncritical borrowing of terms, as in the relating of creative inspiration and 'Lewin's principle of closure' (68). Again, there are rather unfruitful verbal distinctions between external and internal facilitating or inhibiting influences (Ch. II), and between sustaining and energizing motives (pp. 170 ff.).

Three more general criticisms may be raised. First, much in this book depends on the concept of set. It is given definition and mention in several places, but in view of the emphasis on it, a thorough treatment of it might have been appropriate. Second, because of the nature of the style of writing, rather sweeping generalizations are made, as when it is stated that reading and recitation should be employed in a 1 to 4 ratio for efficient learning (239). This statement probably is based on Gates' experiment, which simply does not show this for meaningful material, or for a wide range of conditions. Third, the author points out the close inter-relations of 'mental' and 'physical' fatigue in the first chapter, but thereafter body and mind tend to be independent entities.

At the risk of being unfair to Professor Bills, it may be said that some passages in the book are rather hard to swallow. For example, in a discussion of personality improvement where he is concerned with tastes in reading, music, and art (298) he says:

> The widespread American attitude has long been to reject the best in favor of the popular, on the ground that red-blooded masculinity has no place for refining influences. The defect in popular music is that it depicts such a limited range of moods and thoughts. The lament of the lone cowboy has its place, but a little of it goes a long way. Such music requires little exercise of the nervous system above the spinal cord, and leaves no residue of developed mental muscles, no sense of achieving. . . .

In another instance, having reported that efficiency rises during the work week to a peak on Wednesday, he recommends, with professorial calm,

Thursday is the natural point at which to introduce a half-day rest . . . we herewith advocate that all workers, particularly mental workers, adopt such a plan. Housemaids already observe the Thursday half holiday; why not office workers, industrial workers, business men, and students? (222).

The above criticisms are concerned mainly with form—credit should be given for the effort to bring together in a meaningful way the diversity of studies which pertain to efficiency at work. The book does not leave the reader with the feeling that he has been exposed to a fresh point of view, but this may not be possible until more of the data on efficiency in the work of war are available.

CLAUDE E. BUXTON

University of Iowa

WHITNEY, DAVID D. Family treasures, a popular guide to heredity. Lancaster, Pennsylvania, Jacques Cattell Press, 1943. Pp. 299.

Here is a popular introduction to knowledge about human heredity, written by a zoologist of recognized ability. The author has made numerous important contributions to research in genetics. His book is intentionally factual, it is restricted to presentation of data acceptable to scientists, and it attempts to educate readers against acceptance of prevalent discredited ideas. Emphasizing the inheritance of normal human traits, it gives a clear idea of the meaning of such traits through unusually extensive photographic material.

In the family history technique used in this study of heredity, the compilation consists largely of photographic records showing the inherited characteristics of members of families traced through two to five generations. This material is supplemented by reference to the available published studies, thus placing the photographic data in the general research literature, and lending balance and perspective to the whole.

Evidence concerning about 200 traits is presented. These include such characteristics as obesity, hair color and form, stature, characteristics of the blood, eye-color and form, shape and size of nose, and many details of the hands, feet, trunk, head, and face. Some space is given to unusual characteristics such as albinism, hemophilia, and brachydactyly, but abnormal traits are not emphasized. The photographic method appears to be very effective in giving precise meaning to the trait-names. Probably the weakest chapter in the book is that dealing with temperament. Here the discussion is much less sophisticated than one might expect in view of the extensive work of modern psychologists who have written about temperament and personality. However, in the other, more important parts of the book, Whitney has succeeded in clarifying the meaning of many concepts dealt with in genetics. Such basic ideas as dominance, recessiveness, sex linkage, multiple determiners, and hybridism are effectively exposed.

The book may be used by teachers in developing a general readiness for college courses in genetics. It will interest many individuals who have observed but not understood the significance of traits which run in their families. Anyone interested in the details of personal identification will enjoy the chapters dealing with inherited characteristics of the hands, feet, head, face, eyes, mouth, etc. For example, psychologists interested in the diagnosis of monozygosity in twins will find much useful data. Since details of physique are probably more noticeable to persons who have studied them intensively, it seems likely also that the book could be used to advantage in modern schools for police officers.

UNIVERSITY OF MICHIGAN LIBRARIES

The obviously sketchy nature of much of the data, perhaps emphasized by the method of presentation, calls attention to the need for more research of an organized sort in the field of human heredity. This too is a part of the author's intention. Whitney has been successful in all of his major purposes, for the book is sound, interesting, simple but instructive, and likely to stimulate further data collection. It is written for amateurs; it is likely to lead some of them into professional work in genetics.

HAROLD D. CARTER

University of California

BRECKENRIDGE, MARIAN E., & VINCENT E. LEE. Child development: physical and psychological growth through the school years. Philadelphia: W. B. Saunders, 1943. Pp. ix+592.

For some time there has been a need, which this book helps to supply, for a good textbook on the development of school-age children. In this instance the collaboration of a nutritionist and a psychologist has resulted in a book which deals in some measure with every aspect of the growth of children. The authors say that the book "is designed for professional students in psychology, teacher training, home economics, medicine, nursing and social work as well as for parents." This is a large order. Although it is pretty well filled for the various professional workers, sections of the book are probably too advanced and the language too technical for many parents.

Copious use is made of current research findings: often both sides of controversial issues are presented. Care is taken to state whether or not a point is substantiated by adequate research. The authors quote freely from many sources. The bibliography of 806 titles is an indication of the thoroughness with which they have covered the fields of research in child development. Of necessity the discussion of some topics is brief, but all are well-documented. Suggestions for practical applications are given for each topic discussed, with specific attention often being called to practical methods in the school and in the home. At the end of each chapter are "Questions for Class Study" and references for "Selected Readings".

The organization of the material departs notably from the traditional arrangements in child psychology texts. There is a strong emphasis on the interrelatedness of the various aspects of development. This interrelatedness is given precedence in the chapter arrangements, Chapter I being on "General Principles of Development" and Chapter II on "Interrelatedness of Growth." In this latter chapter considerable use is made of Olson's concepts of organismic age. The following four chapters are concerned with "Influences on Growth." These are grouped into: "Heredity, Endocrines, Illnesses, Emotions," "Nutrition and Routines," "Home, School, Church and Camps," and "Further Community Factors." After all of these discussions of interrelatedness and influences, the authors turn in the next six chapters to the consideration of growth itself, in physical, motor, various mental and intellectual, and personality trends. (Intelligence is treated incidentally under "Sense Perceptions and Judgments.") There are two additional chapters on "Social and Personality Development," the latter discussing in some detail "Moral Judgment and Psycho-sexual Development." The final chapter summarizes the general trends of growth and gives case studies of two typical children.

This organization of the material is well calculated to emphasize the growth of the child as a whole. However it may operate to make the book too difficult

for persons who have not had preliminary work in some part of the field. A person who is trained in any one area of child development should find this book excellent for general orientation to the whole field. It should be a good text for an advanced course, and an excellent source for references to recent research and writings on growing children.

NANCY BAYLEY

University of California

TOMKINS, S. S. (Ed.) Contemporary psychopathology. A source book. Harvard University Press, 1943, Pp. 600.

This book is a collection of 45 condensations of researches in the field of abnormal psychology, for the most part published within the last 5 years. It is divided into four sections: seven papers are on mental diseases in childhood, 10 on psychoneuroses and psychosomatic medicine, 13 on the schizophrenic psychoses, and 15 on experimental psychopathology.

The studies have been well edited. They are presented in sufficient detail to allow the discriminating reader to recognize the obvious merits or limitations of experimental design and procedure. In the individual investigations, the papers have been presented in clear and concise form, are easily understood and may be read rapidly after mastery of the technical terminology frequently used.

At least six of the seven papers on mental diseases in childhood are written from a psychoanalytical point of view. Not all of the papers enjoy such a standard of excellence as to justify this over emphasis. The dangers of coercive and threatening behavior in the development of the infant and young child are emphasized in the well-formulated studies by Huschka. The uses and limitations of release therapy in treatment of the emotional problems of the young infant are clearly presented by Levy.

All reported clinical investigations dealing with psychosomatic medicine are studies carried out at the Chicago Institute for Psychoanalysis. These workers have attempted to relate gastrointestinal, asthmatic and vasomotor disturbances to basic personality types produced by varied predominant conflict-situations. That there is a relationship between certain "organic" disorders and prolonged unconscious emotional conflicts seems to be well established. The identification of a particular disorder with a particular predominant conflict situation is yet to be conclusively demonstrated, and there is general agreement that further research is needed.

An extremely interesting paper on the experimental production and control of peptic ulcer in a man with a gastric fistula is presented by Wolf and Wolf. A statistical study by Dunn points out and attempts to explain the high incidence of gastroduodenal disorders among (English) soldiers during this war. Late in the section a paper by Rosenzweig deals with the importance of implicit common factors in diverse methods of psychotherapy. The discussion of diverse methods comes as a surprise since, in the clinical articles chosen, there was no indication that there was more than one method used successfully in the treatment of psychoneurotic disorders.

In the section on the schizophrenic psychoses the first two papers describe the existence of physiological changes in the bodily economy associated with this psychotic disorder.

Shiply and Kant review and discuss the relative merits of insulin-shock treatments in cases of schizophrenia. The "bleaching of affective life" by frontal

UNIVERSITY OF MICHIGAN LIBRARIES

lobotomy is described by Freeman and Watts. According to the data presented by the authors, the poorest results are obtained in cases of schizophrenia.

Three papers dealing with thinking disorders in schizophrenia are presented. There is general agreement that there is a loss or reduction in the categorical attitude of the patient, and a lack of internal intellectual organization. Differences in the thinking processes of schizophrenic patients, of normal children, and of senile patients are emphasized by Cameron.

Three experimental psychopathological studies on the rat indicate that early frustrations subsequently may produce inhibition of function, hoarding, or regression, as the case may be.

Gantt recommends change of environment for neurotic dogs.

Four papers on hypnosis point out the usefulness of hypnosis as an experimental technique in the analysis of psychopathological phenomena. Papers by Haggard, by Murray, and by Sears present experimental evidence of the influence of emotions and of basic personality mechanisms on varied behavioral tests. The articles by Haggard and by Sears have therapeutic implications. The importance of threat in conflict and frustration and its importance in animal and human experiments is emphasized by Maslow.

The book is described in the introduction as "not a textbook in itself but the auxiliary of a textbook." It will serve this purpose admirably in the field of abnormal psychology and many of the abstracted articles could well be used as collateral reading in courses dealing with emotions, motivations, social psychology or clinical psychology. The value of the book is perhaps enhanced by the fact that the student with little statistical training will have no difficulty in comprehension.

HARRY F. HARLOW

University of Wisconsin

FERNALD, GRACE M. Remedial Techniques in Basic School Subjects. New York: McGraw-Hill Book Company, 1943. Pp. xv+349.

Dr. Fernald is primarily a psychologist both by training and interest. It is the psychological aspects of her techniques which make them valuable aids to the education of children who have difficulty in mastering the basic school subjects. She is an able clinician and appreciates the emotional distress suffered by all children who fail to make normal progress in school. Her major contribution is in demonstrating how this emotional blocking, which so often interferes with further learning once a child has failed, can be overcome.

Though many children fail to learn to read under the conditions of the ordinary classroom in the public schools, genuine and incurable alexia has been found to be very rare—limited to a few cases with demonstrable brain lesions. Dr. Fernald works from the premise that all other children, assuming that they are well and of normal intelligence, can be taught to read by employing a suitable method.

The Fernald method of remedial instruction, which has been briefly described in several previous publications, differs from other methods in drawing heavily upon auditory and kinesthetic functions. Each word to be learned is first written by the teacher and then pronounced naturally for the child. The child traces the written copy of the word with finger contact while saying the word until he feels that he has mastered it. Next he writes the word from memory on a slip of paper which he files for future reference, and writes the word again in the composition for which he needs it. His composition is typed in cor-

rect form by the teacher and constitutes the material for his first reading. No copying is allowed, and no child is ever forced to write a word that he does not feel he knows. This procedure minimizes the practice of errors. Since words of any length or difficulty can be learned by this method, an older child can write about things that interest him and which are in keeping with his mental age.

The Fernald method differs rather markedly from the methods of teaching reading commonly used in the schools of today. The value of this fact in gaining the child's cooperation should not be underestimated. He is asked to try a new way of learning words which many boys who have had difficulty in learning to read have found successful. He sees other boys and girls like himself at the Clinic School obviously making progress. He tries the method and finds that it *is* different. Moreover, his ability to learn by its use is immediately demonstrated to him. Just as in the case of a new medical treatment, the novelty of the approach has tremendous worth in gaining the enthusiastic cooperation of the patient.

The general procedures at the Clinic School are not entirely original but are psychologically sound. Children are first given something to do which they know they can do. The need for reading and writing develops out of this situation. By reference to carefully gathered case history material, it is possible to avoid reinforcement of negative emotional conditioning due to past failures. Once the child has begun to learn, his progress is demonstrated to him as the best possible motivation for further effort. The group instruction available at the Clinic School has an advantage that cannot be matched by the best of individual teaching. Each child enters a smoothly functioning group of children, all of whom have difficulties similar to his own, and all of whom are using the same method he is requested to try. The group situation gives rise to social imitation, provides companionship with other children, and furnishes opportunities for cooperation and competition. Children are not returned to their regular classes until they have enough skill in reading to keep pace with other children of their own ages, and until all subject matter deficiencies have been made up. Unfortunate behavior reactions due to feelings of inadequacy are thus eliminated at their source.

The phenomenal success of the kinesthetic-auditory method described in this book will appear to many to be an overstatement of the facts. It should be remembered that Dr. Fernald does not imply that her method is the only method by which these children could have learned. She does demonstrate with carefully kept records that the children and adults who have come to her, already having failed by use of various other methods, did learn with relative ease and permanence by use of the kinesthetic method. The underlying psychological question has yet to be answered: are there intrinsic differences in sensory functioning and mental organization which account for the failure of certain individuals to learn by some methods? Or, is it largely incidental which method or combination of methods are a part of the total situation in which a child succeeds or fails? More specifically, are there some children for whom auditory, kinesthetic, or primarily visual methods are essential? As a result of her experience, Dr. Fernald is inclined to feel that there are differences in brain functioning which account for non-readers and cases of partial disability. If this is true it should be possible to measure these differences, thus preventing failure by the use of the correct approach to each child in the beginning. On the other hand, if learning is found to depend on the total situation rather than on any particular approach, the general psychological introduction of children

UNIVERSITY OF MICHIGAN LIBRARIES

to the basic school subjects, so ably described in this book, is of the utmost importance.

Several chapters are devoted to applications of the method to mentally retarded children, to the teaching of beginning classes (with careful follow-up study of each child), and to incorporation of the method into the regular procedures of the public schools at any level. Included also are discussions of the application of the method to the teaching of spelling, and of remedial work in arithmetic. The organization of the content could have been improved, thereby eliminating a great deal of unnecessary repetition. The book is attractively set up and generously illustrated with pictures of children engaging in various phases of the remedial instruction, and with photostatic copies of their work. Tables supply details of the progress of individual children. A well selected bibliography of several hundred titles has been appended.

KATHARINE M. MAURER

University of Minnesota

NOTES AND NEWS

The University of Pennsylvania has set up a Personnel Index, a new unit through which alumni now in the armed forces will be assisted in obtaining employment after the war. The service will be conducted on a basis of cooperation between the university and the War Department, the National Roster of Scientific Personnel, and the War Manpower Commission. CLARENCE E. CLEWELL, director of the placement service is director of the service. Assisting him is an advisory committee composed of administrative officers and professors, among whom is ROBERT A. BROTEMARKLE, professor of psychology.

MARGARET D. GORDY, director of placement, Wilson College (Chambersburg, Pa.), has been commissioned to make a six-month survey of tests and their use in occupational adjustment. The study is designed to help the students and alumnae of the college in planning postwar careers.

P. S. DE Q. CABOT has resigned as assistant professor of psychology at Simmons College to accept the position of Director of Personnel Research in the United Drug Co., Inc., Boston. He remains as consultant and Research Associate for the Ella Lyman Cabot Foundation. He has recently completed a confidential war emergency assignment in Washington.

EARLE E. EMME, chairman of psychology and guidance, and director of the downtown division at Morningside College, was recently elected Dean of Dakota Wesleyan University, and professor of education and psychology. Dr. Emme was completing his tenth year at Morningside. He began his new duties at Dakota Wesleyan on March 1.

Wilson College (Chambersburg, Pa.), announces the promotion of Mrs. EDWARD E. ANDERSON to an assistant professorship in psychology.

GOODRICH C. WHITE, president Emory University (Ga.), has been appointed by the Association of American Colleges as one of the association's "three representatives to the American Council on Education."

SETH ARSENIAN, professor of psychology, Springfield (Mass.) College, was granted a leave of absence in December for service as head of the Near Eastern section of the Foreign Broadcast Intelligence Service, FCC, Washington, D. C.

FRANK A. BEACH, chairman and curator of the department of animal behavior at the American Museum of Natural History, addressed members of the faculty and graduate students of the departments of zoology and psychology at Indiana University on February 7. His lecture was entitled "The Neural and Hormonal Factors Involved in Reproductive Behavior."

CHRISTIAN A. RUCKMICK, now connected with the North Western Railroad spoke on Thursday, February 24, at the Milwaukee State Teachers College on the topic "Controlling Your Emotions." Dr. Ruckmick is a member of the Committee on Techniques and of the Committee on Standards of the Chicago Office of Civilian Defense.

The Michael Reese Hospital announces its 1944 course in the Rorschach test to be conducted by S. J. BECK. The emphasis this year will bear especially on the test insofar as it indicates expectation of success in treatment. The dates are June 5–9, inclusive; two sessions daily, two hours each session. Information may be obtained from the Department of Neuropsychiatry, Michael Reese Hospital, 2815 Ellis Avenue, Chicago 16, Illinois.

Psychological Cinema Register. The Psychological Cinema Register, formerly operated by ADELBERT FORD of Lehigh University, has been purchased by The Pennsylvania State College, State College, Pa. The sale of the films listed in the September, 1941, edition of the Register, which was suspended in

UNIVERSITY OF MICHIGAN LIBRARIES

December, 1941, will be resumed immediately. The editorial supervision of the Register will be carried on by the Department of Education and Psychology of The Pennsylvania State College. EDWARD B. VAN ORMER, associate professor of psychology, will serve as editor with the assistance of other members of the department and such other specialists as may seem desirable. The business management of the Register will be in charge of I. C. BOERLIN, supervisor of Audio-Visual Aids Services of the Central Extension Division of The Pennsylvania State College, under whose direction the College maintains a rental library of educational films and other audio-visual aids, a projection service for both campus and extension classes, and a 16 mm. sound motion picture production unit. C. R. CARPENTER, associate professor of psychology at Pennsylvania State and now on leave as a Captain in the Army Air Forces, and chairman of the APA committee on Audio-Visual Aids, will be an advisory editor. In taking over this distribution of instructional films in psychology and related fields, The Pennsylvania State College wishes to be of service in the instructional phases of psychology, psychiatry, and related professions. It plans to make the more popular films available gradually on a rental basis, thus being of service to smaller institutions who cannot afford the convenience of purchasing films. The new editor and manager solicit the continued interest of the psychological, medical, and psychiatric professions in the Register and will welcome suggestions for increasing and improving its service. Communications regarding the submission of new films should be addressed to the *Editor*, EDWARD B. VAN ORMER, The Psychological Cinema Register, Audio-Visual Aids Library, The Pennsylvania State College, State College, Pa. Communications regarding purchase of films and possible rentals should be addressed to the *Manager*, I. C. BOERLIN, at the same address.

Announcement of Senior Internships. Announcement is made of two senior internships (residencies) in psychology open in September, 1944, at the Western State Psychiatric Hospital in affiliation with the Department of Psychology of the University of Pittsburgh under the immediate direction of DR. SAUL ROSENZWEIG. The positions are intended for advanced training in clinical psychology, broadly conceived to include the fields of psychological tests, counselling and experimental psychopathology. Men or women who have a Master's degree in psychology and some clinical experience, and who wish to study for the Doctor's degree are eligible. At the hospital, the candidate will receive advanced training in clinical procedures and carry on research toward the doctoral thesis; at the University, graduate courses for the Doctor's degree will be taken. Full maintenance at the Hospital, including board, room and laundry, is provided in addition to a cash stipend of approximately $75 a month. University courses may be taken at half-rate tuition ($5 per semester credit). Appointments will be made for one year subject to renewal if requirements have been adequately met. Applications should be addressed to DR. GROSVENOR B. PEARSON, Director, Western State Psychiatric Hospital, Pittsburgh 13, Pa.